May 2020

Argument as a Context to Understand Students' Biology Epistemology

Dennis Michael Lee
Clemson University, dennislee286@gmail.com

ARGUMENT AS A CONTEXT TO UNDERSTAND STUDENTS' BIOLOGY EPISTEMOLOGY

A Dissertation
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
Engineering and Science Education

---

by
Dennis Michael Lee
May 2020

Accepted by:
Lisa Benson, Committee Chair
Dylan Dittrich-Reed, Committee Co-Chair
Courtney Faber
Eliza Gallagher

ABSTRACT

Science epistemology, what we know about science and how we know it, is an essential part of scientific literacy. Individuals' science epistemology allows them to comprehend and ascertain the validity of scientific claims, helping biology majors to become better scientists and non-biology majors to influence how biology knowledge informs social policy. As biology education shifts toward active learning and practice-focused approaches, there will be increasing opportunities to discuss what we know in biology and how we know it. For example, Course-based Undergraduate Research Experiences (CUREs) present students with the opportunity to engage in authentic practices that result in biological knowledge production. However, little is known about how students employ their epistemologies while participating in these new kinds of learning environments, specifically how students evaluate, build and justify knowledge.

Recent work on epistemology has revealed that student epistemologies are not coherent structures that can be accessed and examined independent of context. Rather, these developing epistemologies comprise disparate cognitive resources that are activated in response to particular contexts. As such, a phenomenographic approach was employed to investigate the qualitatively different epistemic practices that undergraduate biology majors used in the context of building arguments in a biology CURE.

In this study, twenty undergraduate students were interviewed about the arguments they constructed in a research proposal poster presentation. Analysis of participant poster presentations and interviews revealed that participants approached argument construction

from local (classroom) and global (scientific) perspectives. Additionally, when engaging with scientific information, participants discussed the validity of the information (epistemic cognition) and the processes behind how the information was constructed (epistemic metacognition). Each participant constructed two research questions: one developed in collaboration with instructors, and another constructed solely by the participant. As participants moved between the context of the two research questions, their argument practices shifted between local and global perspectives as well as between epistemic cognition and metacognition. Three distinct practice composites emerged from analysis of participant argument construction across these contexts: Validating, Enculturing, and Transitioning. Validating practices were aimed at matching information and practices to what participants perceived as instructor expectations. Enculturing practices were built around information sharing within a knowledge culture. Transitioning practices were applied when participants perceived differences between instructor-sourced information and information gathered through literature searchers and experiments. As students moved from the periphery of the classroom culture toward the center of this community, they used Validating, Transitioning, and finally, Enculturing practices.

These findings inform instructional practice by outlining contexts in which students discuss scientific knowledge production. Biology educators can create similar contexts to stimulate discussion about what biologists know and how they know it, thereby enhancing student understanding of biology epistemology. Furthermore, these findings support and extend previous research describing epistemology as context-dependent. During their interviews, students also discussed dissonance between their classroom perspectives and their perspectives on how professional science constructs knowledge. These reflections led to student descriptions of their beliefs about biological knowledge. These insights invite future research into how student biology epistemologies develop, and how the culture of the classroom contributes to the development of these epistemologies.

DEDICATION

To my grandparents, Florence and Sing-Man Lee. I wish you were here to share this moment with me.

ACKNOWLEDGMENTS

First, I would like to acknowledge my family for all the support they gave me throughout this journey. Thank you, Grandpa, for pushing me to pursue my Ph.D., and for providing the financial means for me to do so. Thank you, Grandma for showing me the importance of education. Thank you Dad, for sharing with me your own experience as a Ph.D. student, and for the guidance you provided. Thank you, Mom for all the emotional support, and the occasional push I needed to continue my work. Thank you Jeff and Becky for being my siblings. Your constant teasing as children taught me that I could endure anything. Thank you Amy, for enduring my stress-induced mood swings and my passionate ravings about epistemology. I feel like I can be myself around you, and that is a gift I can never thank you enough for. And thank you to my children, Ethan and Christa, who through our discussions helped me more deeply understand the development of epistemology in children.

Second, I would like to thank my committee, who provided the insight to make this pile of words into a dissertation. Courtney, Dylan, and Eliza, thank you for reading, critiquing, and making suggestions on my project. Thank you for all the positive reinforcement, and telling me not necessarily what I wanted to hear, but what I needed to hear. Thank you, Lisa for being an excellent advisor. As your student, I felt wanted, confident, and free to pursue my own interests and passions. You were always there to steer me back onto the path when I lost my way, and you were especially adept at using the words “future work” when my ideas got too big. You are the kind of advisor I aspire to be.

I would also like to thank my fellow ESED students, both past and present. I treasure all the deep conversations I have had about education research as we grew together. I learned so much from you, and I hope some of you have learned something from me.

Finally, I would like to thank my participants; without you, this study would not have been.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

1.1. Educating Future Scientists

Young biologists are entering a field brimming with new innovations ranging from the gene editing capabilities of clustered regularly interspaced short palindromic repeats (CRISPR) (Horvath & Barrangou, 2010) to the development of biologically inspired robotics (Chen et al., 2017). These innovations and resulting discoveries often occur at the intersection of multiple disciplines, requiring scientists to view information through a variety of lenses. Moreover, socio-scientific issues such as genetic manipulation and climate change have implications that reach beyond the context of science research into the realm of society. Students need to be able to leverage previously acquired knowledge to create innovations that can address these complex socio-scientific issues. It is therefore our responsibility as science educators to equip our students with an understanding of how knowledge is constructed in science, and the cognitive tools to determine what counts as reliable knowledge in our discipline.

Reviews on how well college prepares students for society and the workforce are mixed. Business and hiring managers believe that it is essential for future employees to complete a college degree. However, while a majority (60%) of hiring managers believe that recent graduates have the skills and knowledge to succeed in an entry-level position, only 25% of hiring managers believe that recent graduates have the skills and knowledge to be promoted to more advanced positions (Hart Research Associates, 2018). Most hiring managers (65%) feel that colleges should make improvements to ensure that students graduate with important skills such as critical thinking, analytical reasoning and ability to solve complex problems (Hart Research Associates, 2018). Students' abilities to solve these types of complex, multidisciplinary problems are supported when they have an understanding of what counts as knowledge, and of the methods used to validate knowledge in their discipline (Hallman, 2017; Lombardi, Nussbaum, & Sinatra, 2016). In other words, students need to apply context-appropriate epistemic practices.

1.1.1. Epistemology and Scientific Literacy

Understanding the ways in which science accumulates, justifies, evaluates and generates knowledge (science epistemology) is an essential component of scientific literacy (Hallman, 2017). This is argued throughout the literature, including the utilitarian argument that science epistemology establishes a set of processes that can be applied to contexts that are solely scientific as well as mixed contexts such as those presented as socio-scientific issues (Driver, Learch, Millar, & Scott, 1996). This utilitarian argument connects to a democratic argument, as these processes can be used by the general public to make decisions about socio-scientific issues that affect the general public such as health, quality of life and the environment (Driver et al., 1996; Hallman, 2017). Addressing socio-scientific problems such as climate change is made even more complex with the explosion of information available through the Internet; while individuals can access information, they need to recognize pseudoscience and evaluate scientific claims using sound evidence. Individuals use epistemic practices to evaluate these scientific and pseudoscientific claims

(Kelly, 2008), and individuals decide when it is appropriate to use specific epistemic practices based on the context and their own science epistemology (Chinn, Rinehart, & Buckland, 2014), often tacitly (Hofer, 2004). Therefore, developing a science epistemology allows individuals to build skills that executives and hiring managers look for in recent graduates, for example, appropriately gathering, organizing and evaluating scientific information (Hart Research Associates, 2018).

1.1.2 Epistemology in Biology

Modern biology epistemology developed from a split in the biology community between experimental biologists and naturalists (Magnus, 2000). Experimental biologists felt that controlled experiments provided objective and therefore reliable knowledge. On the other hand, naturalists felt that multidisciplinary and varied lines of evidence were most important in the support of valid claims, an epistemology described as consilience of inductions (Whewell, 1847). In the late 19th and early 20th centuries, the biology field shifted its epistemology towards experimental control because it was thought that this way of thinking was more objective, and aligned with the rigor of the scientific disciplines of physics and chemistry (Magnus, 2000).

The narrowing of the lines of evidence considered valid in biology is reflected in the recommendations for biology education to equip students with the skills to generate and test hypotheses through experimentation (Brewer & Smith, 2011; States, 2013). However, relying only on experimental processes may lead to misleading results. Magnus (2000) explains in his case study of a debate over the importance of isolation in evolution that one scientist's over-reliance on experimental process led to an incorrect theory that evolution needed only mutation to form new species. In fact, naturalists that relied on consilience had already seen the limitations of this theory because they used a larger pool of evidence that included information experimentalists felt were too subjective to be reliable (Magnus, 2000). Naturalists' contributions led to the modern evolutionary theory that new species formation is influenced by, among other things, a mixture between mutation and geographical isolation.

This is not to say that experimental epistemology is flawed; on the contrary, it is an effective way of thinking that is prevalent in science. It was the variation in lines of evidence, which included experimental data, that allowed naturalists to see the flaws that pure experimentalists could not. Consequently, there have been calls to move toward an epistemology of biological pluralism, where biologists evaluate knowledge through a multitude of epistemic lenses (Longino, 2000). To move toward this approach, it is important to understand the affordances and limitations of the epistemologies applied to gather, justify, and construct knowledge. Imparting this understanding to our students facilitates their application of appropriate epistemologies in varying contexts.

1.1.3 Education in Biology

Biology education has seen a shift from teaching students *what they need to know* to teaching students *what they need to do* (Duschl, 2008). While contemporary life sciences curricula focus on both concepts and processes related to the discipline of biology (Bybee, 2012), the epistemological underpinnings of biological knowledge remain tacit. For example, the Next Generation Science Standards (NGSS) mandate that students in science courses "develop useable knowledge that can be applied across the science disciplines" (NGSS Lead States, 2013, p. 75), but they do not comment on what counts as useable knowledge, or how that knowledge could be gathered, justified, or constructed. Focusing instruction on memorization of core biology concepts and processes may lead to learning content and skills without understanding why or in what context those concepts and skills should be applied (Berland et al., 2016). This is problematic since the core aim of science is to describe natural phenomena through the acquisition, justification, and generation of knowledge (Longino, 2002). Performing the actions of science without the core aim of acquiring, justifying or generating knowledge cannot be considered science practice.

1.2. Purpose

Few research studies are performed solely for the sake of knowledge; there is often a practical issue that is the driving force behind the research. The driving force behind my research is to improve biology education by teaching students to think like a biologist. In teaching a student to think like a biologist, an instructor may weigh the correctness of a student's way of thinking, or may try to understand the student's thinking process. An effective instructor should endeavor to do both to measure students' progress toward thinking like a biologist (Marton & Booth, 1997). This project focuses on understanding students' thinking processes because to understand why students think in different ways, we must first understand the different ways in which students are thinking.

There are few studies that examine how teaching students biology through scientific practice influences their science epistemology development. This study is the first step in achieving the overarching goal of examining how student epistemologies develop while participating in biology curricula, and how these curricula can be leveraged to facilitate the development of student epistemologies that are productive in their chosen disciplines, careers and everyday lives. The first step in this inquiry, and the purpose of this study, is to examine how students think about knowledge in a biology classroom. One difficulty in studying epistemology is that an individual's thoughts about knowledge are often tacit, which necessitates the study of epistemology in the context of a task where the subject is engaged in a knowledge process (Hofer, 2004). *Thus, the overarching goal of this project is to examine the qualitatively different ways in which biology students use epistemic practices to build arguments.*

1.3. Project Overview

In this study, I sought to understand the myriad ways in which biology students use epistemic practices to build arguments. To this end I designed this study with phenomenography in mind, organizing differing ways of experiencing a phenomenon into categories of description, and describe the relationship between these categories in an outcome space (Åkerlind, 2012; Bowden, 2005; Marton, 1986). Categories of description organize distinctly different ways of understanding a phenomenon into groups, and the outcome space that describes the logical relationship between each category of description (Marton & Booth, 1997). To construct categories of description (Chapter 5) and an outcome space (Chapter 7), I collected student research proposals, identified student arguments within the text, and asked students about how they constructed these arguments in semi-structured interviews. I worked with an analysis team to identify passages in the interview transcript that were epistemic, then open-coded the identified passages. Following the first cycle of analysis, we found that the categories of description only described student epistemic practices at a surface level. As a result, we conducted two more cycles of analysis by first analyzing interviews question by question across all participants, and then by analyzing each participant across all interview questions. Through these cycles of analysis we were able to construct categories of description and an analytic framework (Chapter 6) to contextualize how participants used their epistemic practices to build the arguments in their research proposals..

While the categories of description illustrated the qualitatively different ways in which biology students constructed arguments, they did not sufficiently show how these practices interacted as a whole. To illustrate how individuals used these collections of practices in different contexts within the class, I organized the categories of description into three composites, and present three exemplar cases (Yin, 2018) that describe how each participant constructed their arguments and justified their claims (Chapter 7). These exemplar cases follow participants through the arguments they make and give us a sense for why participants chose to use particular epistemic practices in different contexts. Since this project was not longitudinal, I am unable to explain how participants developed these epistemic practices.

1.4. Study Limitations

There were several limitations for this study. First, the study population was selected from an honors biology curriculum, so I caution readers from considering participants in this study as general student population. Since students were required to apply to the program and meet requirements such as minimum GPA, prerequisite coursework, or advanced placement, I remind readers that results from this study should only be compared to similar contexts and should not be generalized. Second, data for this study were collected from a single institution. While participant recruitment was designed to achieve maximum variation, it is nonetheless limited to the amount of variation available within the range of students enrolled in the curriculum studied at this institution. Third, the data for this study were collected all at once within a short time period to ensure research quality of the phenomenography. As a result, the data presented represent a snapshot of student epistemologies assembled from participant recollections and artifacts. Since this was not a longitudinal study, no conclusions can be drawn about the development of student epistemologies or epistemic practices.

1.5. Positionality Statement

I am a former doctoral student in molecular biology and have a Master's degree in microbiology, which affects my beliefs about what counts as scientific knowledge. Before enrolling in my current doctoral program, I felt that experimental data, particularly quantitative experimental data, was the most valid form of knowledge in science. Through my experience learning about education research methods, I have come to realize that there are contexts where it is appropriate to apply quantitative experimental methods, and there are other contexts where it is appropriate to apply qualitative methods. As a result, I feel that knowledge should be evaluated from multiple epistemic perspectives. These ideas allow me to consider a diversity in types of evidence, aligning my epistemology with consilience.

Evaluating knowledge from multiple epistemic perspectives brings with it the problem of underdetermination, the notion that multiple valid interpretations can be drawn from identical data. I reconcile these contradictions through my belief that knowledge is uncertain, and that truth is present – and different – for each individual. This uncertainty is tempered by the strength and volume of evidence supporting each claim. My thoughts about knowledge color my interpretation of the data and the conclusions that I draw from this study.

Furthermore, my experience as an introductory biology instructor influences how I view students, and how I interpret their actions. Through my teaching experience, I have come to believe that many students are grade-focused, and acquire information for the purpose of fulfilling criteria set by classroom assessments. This belief colored my interpretation of student actions in this research, and I endeavored to mitigate this bias in my analysis. This belief also drives me to improve biology education by extending student goals beyond "getting the grade" to constructing and applying knowledge to solve the world's complex issues.

2. LITERATURE REVIEW

Researchers have adopted two main approaches to identify goals for teaching and learning epistemology: by studying lay or expert epistemologies (Barzilai & Chinn, 2017). However, few studies have investigated the epistemologies students employ to construct knowledge in learning environments such as undergraduate research experiences (UREs). Understanding specifically how *students* conceptualize knowledge and knowing in the science classroom is important since it has been found that students' formal epistemologies (their beliefs about how knowledge is gathered, justified and constructed in professional science) may be different from their practical epistemologies (how they gather, justify or construct knowledge in practice) (Sandoval, 2005). Moreover, researchers have found that science epistemologies differ across scientific disciplines (Knorr-Cetina, 1999; Samarapungavan, Westby, & Bodner, 2006). Since STEM graduates will be expected to meet challenges that encompass multiple disciplines, we should not require students to pursue a prescribed epistemology *of* science, rather, we should encourage students to engage in a variety of practices that help to achieve knowledge: an epistemology *for* science (Russ, 2014). Given that students' practical epistemologies differ from their formal epistemologies, this study will investigate student epistemologies by examining student arguments in the context of an inquiry biology course.

In the following sections I will articulate my motivations for this study by briefly summarizing the development of epistemological theory from lay and expert epistemologies. Following this summary, I will address the importance of context to epistemology, and describe how the prevalent use of surveys in biology education research on epistemology may limit our understanding of student biology epistemology. Next, I will connect the practice of argumentation to the study of knowledge. Finally, I will end this chapter with a description of my theoretical frameworks, and the ways in which I will operationalize them in my data collection and analysis.

2.1. Theories Based on Lay Epistemologies

An influential qualitative study by Perry (1970) established the roots of epistemology from a developmental perspective through the longitudinal study of 85 male and two female Harvard University students. Perry conducted unstructured interviews, asking participants "What would you like to say that stood out to you during the year?" along with follow-up questions to develop a framework consisting of nine "positions." Later, these positions were grouped into four categories representing a developmental spectrum: Dualism, Multiplicity, Relativism, and Commitment Within Relativism (Hofer & Pintrich, 1997; Perry, 1970). Individuals characterized as dualists perceive information as either correct or incorrect, with little appreciation for nuance or context. As subjects move into the multiplicity position, they see that authorities may disagree, pushing the subjects to believe that all answers are valid and based on individual opinion. Relativists consider the context while assessing the validity of knowledge. Finally, the position of commitment within relativism is achieved when an individual begins to affirm decisions through a focus on responsibility and values they hold. Subsequent studies expanded on Perry's scheme, producing other theoretical frameworks such as Women's Ways of

Knowing (Belenky, Clinchy, Goldberger, & Tarule, 1983), Epistemological Reflection (Baxter Magolda, 1992), Reflective Judgement (King & Kitchener, 1994), and Argumentative Reasoning (Kuhn, 1992). While each framework expanded on Perry's scheme, connections to his framework were evident in the qualitative methodologies and the descriptions of the developmental stages. Critics of these developmental descriptions of epistemology found that subjects would not meet all criteria within each developmental stage in sequence (Hofer & Pintrich, 1997; Schommer, 1990), and subsequently proposed that epistemologies are comprised of dimensions that act independently of each other (Schommer, 1990)

Thus, in the early 1990's, psychological studies on epistemology began to shift from strict developmental stages to more fluid descriptions of epistemic beliefs. Schommer (1990) developed a 63-item quantitative survey that measured the epistemic beliefs of undergraduate students, and found that students held distinct epistemic beliefs that affected their comprehension and learning. Through a thorough review of relevant literature, Hofer and Pintrich (1997) built on Schommer's and others' work, proposing four dimensions of epistemology: certainty (knowledge is absolute ←→ tentative), simplicity (knowledge is isolated ←→ interrelated, sources (knowledge comes from authority ←→ derived from reason), and justification of knowledge (knowledge is fact and needs on justification ←→ knowledge is constructed and is constantly re-evaluated). Each dimension exists as a continuum between two extremes. While these theories moved away from strict developmental stages, they were still predicated on the assumption that individuals hold beliefs along a spectrum for each dimension, together comprising a set of epistemic beliefs that remain stable across contexts (Hofer & Pintrich, 1997).

About the same time that Hofer and Pintrich (1997) were formulating their epistemological dimensions framework, Kruglanski (1990) developed the lay-epistemic theory that outlined individuals' information processing through hypothesis generation and validation. Kruglanski identified the epistemic motivation for hypothesis generation and validation as the individuals' need for specific or non-specific closure. From this work an instrument was developed and validated for measuring need for closure differences among individuals, and revealed five factors: Preference for order, Preference for predictability, Decisiveness, Discomfort with ambiguity, and Closed-mindedness (Webster & Kruglanski, 1994). While the epistemic dimensions framework focused on an individual's views about knowledge (Hofer & Pintrich, 1997), lay-epistemic theory attended to the motivations behind information processing (Kruglanski, 1990). Nevertheless, both Hofer and Pintrich's epistemic dimensions framework and Kruglanski's lay-epistemic theory present a generalized vision of a lay individual's epistemology that is consistent across contexts. However, some have argued that epistemologies do not comprise stable, coherent constructs like beliefs, rather they propose that epistemologies are formed from disparate, fine-grained cognitive resources that are applied in particular contexts (Louca, Elby, Hammer, & Kagey, 2004). The contextual nature of epistemology will be discussed later in this chapter. Furthermore, individuals' discussions about their lay epistemologies do not represent their optimal level of thinking (King & Kitchener, 1994) and thus may not represent the most productive means of pursuing knowledge. Researchers have therefore pursued an alternative avenue to identify productive epistemologies for scientific thinking: expert epistemologies. Characterizing the epistemology of scientific experts ensures that the described epistemologies are productive for knowledge production because of the success experts have had in contributing knowledge to their field (Barzilai & Chinn, 2017).

2.2. Theories Based on Expert Epistemologies

Epistemology research in science has focused its lens specifically on expert ways of knowing in science, one of the most well-known examples being the Nature of Science (NoS). NoS originates from discussions on how to improve science education, drawing from the history and philosophy of science (Elby *et al*., 2016). While there is still debate on what constitutes NoS, there is agreement on a general set of NoS learning objectives based on what is accessible to K-12 students (Lederman, 2007). There are seven aspects to NoS: the ability to 1) understand the differences between observation and inference, 2) distinguish between laws and theories, 3) understand that science works with model systems, 4) understand that science is subjective, 5) understand that science is performed in the context of culture, 6) understand that scientific knowledge is not certain, and 7) understand that NoS, while deeply intertwined with scientific practice, is a separate concept (Lederman, 2007). These outcomes have been used as a standard to assess both teacher and student understanding of NoS (Bell & Lederman, 2003; Bell, Matkins, & Gansneder, 2011; Lederman, Abd-El-Khalick, Bell, & Schwartz, 2002; Lederman, 1992). The limitations of using expert epistemologies to define the goals of epistemic education have to do with context. First, while there is agreement on the “Lederman 7” described above, there is little agreement about any other aspects of the ideal epistemic practices of science (Barzilai & Chinn, 2017). Moreover, students may view contexts differently from experts (Sandoval, 2005), and therefore may recognize a context as non-scientific when an expert would recognize the same context as scientific. As a result, students may apply less productive epistemic practices even if they achieve the seven outcomes described by Lederman.

2.3. The Role of Context in Epistemology

While NoS and other epistemic belief frameworks suggest that individual epistemologies are coherent cognitive structures that are stable across contexts, some researchers have proposed that an individual’s epistemology is sensitive to context (Clark a. Chinn, Buckland, & Samarapungavan, 2011; Louca et al., 2004; W. Sandoval, 2014; Watkins & Elby, 2013). As such, students’ reported epistemic beliefs may shift depending on the context. For example, in one study, a physics student activated the epistemic resource “knowledge as propagated stuff” when studying for an exam, but activated the epistemic resource “knowledge as constructed” when tutoring a peer (Elby & Hammer, 2010). In a study comparing science majors and non-science majors, researchers found that psychology majors would use general search engines when looking for information outside their discipline even though they believed that specialized psychology search engines should be used when searching within the discipline of psychology (Hofer, 2004). These instances have led some researchers to believe that epistemology should include not only generalized epistemic beliefs, but also the contextualized epistemic practices individuals use to process information (Sinatra, 2016).

2.4. Research on Epistemology in Biology Education Using Surveys

The emerging epistemology research in biology education has focused on assessing the effectiveness of teaching interventions using surveys. Student responses on surveys following the implementation of an active learning intervention in a large classroom showed that students saw knowledge in biology as a collection of facts transferred from professor to student (Walker, Cotner, Baepler, & Decker, 2008). Supporting this finding are survey results that indicated student perceptions of science epistemology, especially those related to problem solving and memorization, shifted towards a more novice-like perspective during an introductory biology class between the beginning and end of the semester (Semsar, Knight, Birol, & Smith, 2011). However, not all assessments of science epistemology have indicated a shift toward novice-like views. Survey results from community college students (Kenyon, Onorato, Gottesman, Hoque, & Hoskins, 2016), first year students (Gottesman & Hoskins, 2013), and advanced students in four year colleges (Hoskins, Lopatto, & Stevens, 2011) exhibited enhanced understanding of science epistemology after exposure to a pedagogy involving analysis of scientific literature called "Consider, Read, Elucidate hypotheses, Analyze and interpret data, Think of the next Experiment (CREATE)" (Gottesman & Hoskins, 2013; Hoskins et al., 2011; Kenyon et al., 2016). These survey results present a view of biology students' generalized epistemologies, however the differences in student understanding of science epistemology may also reflect limitations that result from using decontextualized quantitative items to measure a contextual entity like epistemology. For example, one item in the Epistemological Beliefs Scale asks “The only thing you can be sure of is that nothing is sure” (Schommer-Aikins, Duell, & Hutter, 2014). An individual may answer on one end of the scale when thinking about science knowledge, but may answer at the other end of the scale when thinking about religious knowledge. Indeed, some have questioned the validity of certain items in the Epistemological Beliefs Scale (Hofer & Pintrich, 1997), while others have found that items load into different factors when used on a different student population (Qian & Alvermann, 1995).

Studying classrooms that integrate scientific practices into the curriculum and modifying survey items to reflect scientific practice use in classrooms may be a way to circumvent some limitations of surveys, but studying practices in this context presents another set of limitations. Practices and beliefs that are meaningful to science may not be meaningful to the classroom community (Berland et al., 2016). For example, viewing knowledge as transmitted from an authority may not be a productive science practice, but it is quite productive in a classroom where tests focus on fact memorization. Thus surveys, which provide limited opportunity for elaboration, may miss the nuance and context required to fully understand a student’s perceptions of biology epistemology (Watkins & Elby, 2013).

Qualitative methods, such as interviews and focus groups offer opportunities for subjects to elaborate on the nuances and context with respect to their epistemologies. However, directly asking participants to describe their epistemologies assumes that individual epistemologies are concrete, and can be accessed readily by the participant. In reality, individual epistemologies are often tacit (Hofer, 2004), and are activated through actions that require evaluation or generation of new knowledge such as building an argument (Hofer, 2004; Sandoval, 2005).

2.5. Connecting Epistemology to Argument

"No matter how implicit or unaware of it the subject is, some sort of epistemological theory regarding the process of knowing underlies a subject's argumentative reasoning" (Kuhn, 1991, p.172). Using this assumption, Kuhn formed the foundation of her Argumentative Reasoning framework, which proposed Perry-esque developmental stages (Kuhn, 1992). Kuhn (1991) describes three stages: absolutists, who believe that it is possible for experts to know the cause of phenomena in certain terms; multiplists, who believe that all viewpoints are valid; and evaluatists, who believe multiple viewpoints can be valid and evaluate them for their merit.

Not only is Kuhn's Argumentative Reasoning framework useful for characterizing how individuals argue, its development speaks to how effective argumentation can be as a means for studying epistemology. Like the other Perry-esque developmental models, the Argumentative Reasoning framework presents a generalized view of an individual's epistemology. However, argumentative reasoning has been found to be related to education level (Kuhn, 1992). This suggests that argumentative reasoning is contextual: an individual's education level can differ between disciplines, thus affecting their argumentative reasoning.

In her study, Kuhn found that individuals with different beliefs about knowledge approached argument in distinct ways. Absolutists see little need for argument since they believe knowledge is certain. Multiplists also see little need for argument, but because they see all viewpoints as valid, it would be futile to try to convince someone else of their viewpoint (Kuhn, 1991). On the other hand, Kuhn proposes that evaluatists view argument as an important process of knowing (Kuhn, 1991), and have been found to provide more counter-arguments (Kuhn, 1991) and science-centered data in their arguments (Nussbaum, et al., 2008). These observations suggest that teaching students the practice and importance of argumentation could shift them from absolutist and multiplist stages to the evaluative stage of argumentative reasoning. As a result, some have proposed that instruction in argumentation is an important part of learning (Duschl & Osborne, 2002; Kuhn, 1992; Nussbaum et al., 2008). Indeed, when students were asked to argue alternative positions on a physics problem, they were subsequently more likely to use the correct reasoning on a different physics problem (Nussbaum & Sinatra, 2003). Thus, it is important to study student epistemologies in academic environments that provide opportunities for argumentation, such as undergraduate research experiences.

2.6. Course-based Undergraduate Research Experiences (CURES) as a Context to Study Epistemology

As biology education shifts from teaching students *what they need to know* to teaching students *what they need to do* in science (Duschl, 2008), it presents more opportunities to study student epistemologies as they engage in authentic knowledge production. Undergraduate research experiences (UREs) are an effective way for students to learn the processes of science (Hunter, Laursen, & Seymour, 2007; Lopatto, 2004, 2007; Thiry, Weston, Laursen, & Hunter, 2012; Thiry, Laursen, & Hunter, 2005). However, access to UREs can be problematic because of limited number of available positions and selectivity based on grades and letters of reference from faculty (Linn, Palmer, Baranger, Gerard, & Stone, 2015). CUREs provide an alternative to the URE apprenticeship model through inquiry-based investigations presented as part of the curriculum. As such, students who want to participate simply sign up for the class, as opposed to having to court faculty members by initiating direct contact with principal investigators. CUREs were recently defined as student or instructor driven projects pursued during class that expose students to scientific practices including iteration, with unknown outcomes, and whose impacts extend beyond the course (Auchincloss et al., 2014). Many of the studies on CUREs examine the practical aspects of implementing a CURE, focusing on description of the CURE itself. Assessment of CUREs show a multitude of gains that are similar to students who participate in research apprenticeship type UREs. Participation in CUREs leads to positive outcomes such as increased graduation rates from STEM programs (Rodenbusch et al., 2016), confidence in using lab skills (Rowland, Pedwell, Lawrie, Lovie-Toon, & Hung, 2016), and understanding of NoS (Russell & Weaver, 2011). However, the gains in understanding epistemology were not as robust in CUREs as they were in traditional UREs (Russell & Weaver, 2011). Additionally, researchers have suggested that CUREs help make science more inclusive by removing barriers to participating in scientific research (Bangera & Brownell, 2014). Another study revealed that students' conceptions of what it means to think like a scientist shifted from an initial description of character traits to a description grounded in scientific practices (Brownell et al., 2015). Brownell *et al.* (2015) interpreted this change as a shift toward a more expert-like conception of how scientists think. Interpretation of student exam responses indicated that students improved their data analysis and interpretation skills during their participation in a CURE (Brownell et al., 2015). Olimpo *et al.* (2016) employed a comparison approach between students in a CURE and non-CURE introductory cell and molecular biology course and found that students in the CURE course were more successful in clarifying their career paths, acquiring knowledge in the domain, and felt more self-efficacy and determination.

While these gains are significant, the full list of outcomes likely to be achieved in CUREs remain unclear (Auchincloss et al., 2014). Furthermore, the assessments have not adequately addressed student understanding of the scientific practices or underlying epistemic beliefs of scientific processes in biology. Even though some have called for explicit instruction on the nature of science and science epistemology (Bell & Lederman, 2003), much of biological instruction currently operates under the assumption that practicing biology in a particular way will help students acquire a tacit understanding of biological epistemology. However, while students may believe that scientific knowledge is subjective (a belief that is consistent with NoS), this belief may not be sufficient to help a student develop strategies to engage in productive knowledge construction in science (Barzilai & Chinn, 2017). This idea has led some researchers to suggest that student epistemologies should be studied through their epistemic practices (Hofer, 2004; Sandoval, 2005), which Kelly (2008) defines as the actions students take to find, justify, and construct knowledge. Consequently, biology education's focus on scientific processes provides an opportunity to study students' biology epistemology by examining their epistemic practices. This is not to say that students' beliefs about knowledge are not an important aspect of student epistemologies. On the contrary, students' ideals about what counts as knowledge plays a central role in their selection of epistemic practices. Taken together, students' ideas about knowledge and their epistemic practices comprise epistemic cognition, a construct that I will describe in the following sections of this chapter.

2.7. Theoretical Frameworks

2.7.1. The AIR Model of Epistemic Cognition

It is important to understand that while epistemic beliefs and practices are often conflated, they are distinct. Epistemic beliefs represent *what* an individual thinks about knowledge, while epistemic cognition represents *how* an individual processes information. Accordingly, Chinn *et al.* (2014) proposed a model that views epistemic cognition not as formal beliefs, but as a collection of epistemic practices that are used for the purposes of reasoning such as those used in argumentation. The AIR model of epistemic cognition consists of three components: epistemic *A*ims and values, epistemic *I*deals, and *R*eliable processes to achieve these aims. Epistemic aims and values focus on the individual's aims and the values they place on those particular aims (Chinn, Rinehart, & Buckland, 2014). For example, a student may aim for understanding material for an exam, but may not place much value on that material because the class is not part of their required curriculum. In this case, the student's aim is epistemic but they may not be motivated to pursue this aim.

On the other hand, a politician may seek information to win an argument so they can seek higher office. In this case, the politician highly values the aim and is therefore motivated to fulfill it.

The two examples provided above also influence epistemic ideals, the second part of the AIR model. Epistemic ideals are the guidelines individuals use to determine whether a statement, claim or idea counts as knowledge (Chinn et al., 2014). In the case of the student, understanding material for the exam could drive them to seek justified truth (at least from the professor's point of view). Doing so would hopefully result in a higher grade on the exam for the student. However, the politician's ideals may be that the information gathered agrees with the preconceived notions of their party, eschewing the importance of justified truth. In each case, the epistemic ideal aligns with the context.

These epistemic ideals influence what kinds of reliable processes individuals use to gather and evaluate information (Chinn et al., 2014). These processes may be reliable in certain contexts but not in others. For example, the student may ask their professor about the information provided in class. This is reliable as long as the student is taking the professor's class. However, if the student continues to rely on this professor's expertise while taking a different class, the process becomes less reliable.

The AIR model for epistemic cognition is beneficial to understanding the epistemic practices students use to build and evaluate arguments for several reasons. First, each part of the AIR model takes into account the importance of context in epistemology. Each argument is built and evaluated under certain contexts. The aims and values are influenced by the context, which helps establish the epistemic ideals, which, in turn, help the individual choose the reliable processes to use in their epistemic cognition (Chinn & Rinehart, 2016). Other theories, such as the dimensions of epistemology (Hofer & Pintrich, 1997) provide a generalized view of epistemic beliefs, and do not take context into account. These epistemic beliefs have been observed to change moment-to-moment and between contexts (Andrew Elby & Hammer, 2001). Second, epistemic beliefs are typically tacit and difficult to determine (Hofer, 2004). Therefore, some researchers have suggested that epistemic practices and processes be studied (Sinatra, 2016). The Reliable processes in the AIR model allow us to interpret these processes through its interpretive lens. Since building and evaluating arguments are processes, this model fits well with the research question.

The major methodological challenge to operationalizing this theory is that epistemic cognition happens in large part inside the individual's mind. Therefore, information about epistemic cognition is often tacit (Hofer, 2004). Furthermore, epistemic cognition often changes as the context changes (Chinn et al., 2014; Elby & Hammer, 2001). To operationalize this theory, we need to observe contexts where students are using the processes of epistemic cognition such as while building and evaluating arguments.

2.7.2. Epistemic Thinking

A natural progression of broadening the definition of epistemic cognition is to integrate epistemic cognition with metacognition. In his seminal publication, Flavell (1979) split metacognition into four constructs: metacognitive knowledge, experience, goals, and actions. Metacognitive knowledge is what an individual knows about their own or others' knowledge, metacognitive experiences are ways that an individual experiences a cognitive activity, metacognitive goals refer to any objectives of a cognitive activity, and metacognitive actions refer to how an individual accomplishes these goals (Flavell, 1979). Looking back at the AIR model of epistemic cognition, we can see relationships between epistemic aims, ideals, and reliable processes and the metacognitive constructs described above. For example, some epistemic aims may be considered metacognitive goals if the objective of cognitive activity is knowledge justification or construction. For example, a student's aim could be to justify their experimental conclusions by citing theory from peer-reviewed literature, which is a metacognitive goal. Moreover, a student's epistemic ideals may be derived from what they know about how scientists construct knowledge, which is an example of metacognitive knowledge.

There are several frameworks that integrate metacognition with epistemology (Hofer, 2004; King & Kitchener, 1994; Kuhn, 1999); I chose to use the epistemic thinking framework (Barzilai & Zohar, 2014, 2016) in this study because it specifically integrates the AIR model of epistemic cognition described above with epistemic metacognition

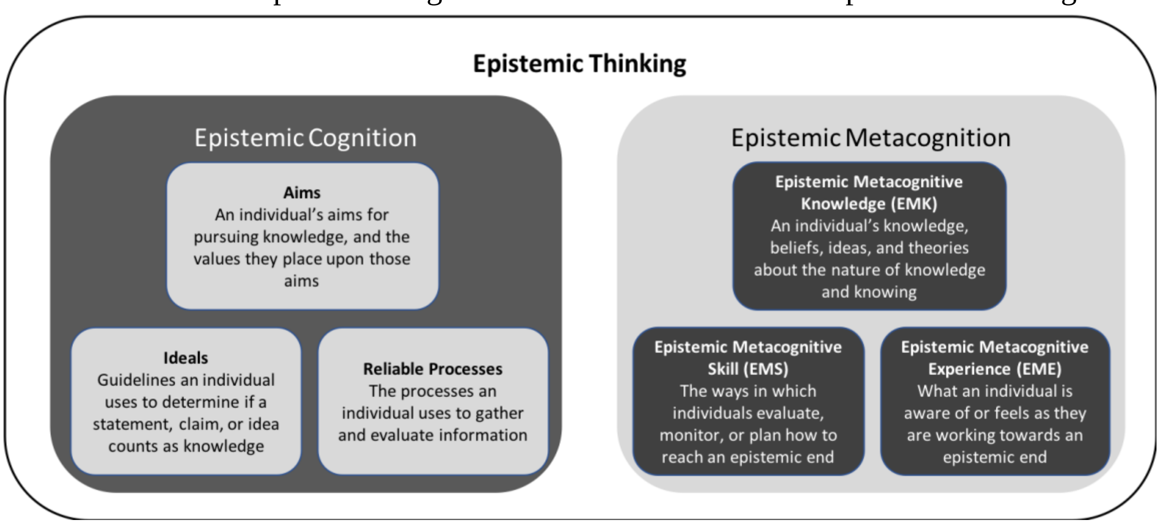

Figure 2.1. The Epistemic Thinking Framework as described by Barzilai and Zohar (2014, 2016).

(Figure 2.1). Barzilai & Zohar (2014, 2016) define epistemic metacognition as how an individual thinks about epistemic cognition. Epistemic metacognition is split into three subcategories: epistemic metacognitive knowledge (EMK), epistemic metacognitive skills (EMS), and epistemic metacognitive experiences (EME) (Barzilai & Zohar, 2014). EMK refers to an individual's thinking about their own or others' knowledge, as well as strategies that disciplines employ to validate or construct knowledge (Barzilai & Zohar, 2016). EMS refers to the ways an individual monitors, plans, or evaluates how to reach an epistemic end (Barzilai & Zohar, 2016). Finally, EME refers to what an individual is aware of or feels while pursuing an epistemic end (Barzilai & Zohar, 2016). The epistemic thinking framework allows me to broaden my analysis to include instances where students are pondering their own or others' epistemic cognition.

One step to operationalize the epistemic thinking framework in my analysis is to determine what counts as epistemic cognition, and what counts as epistemic metacognition. Passages that focus on the validity, truth, or accuracy of the *information itself* would be considered epistemic cognition. On the other hand, passages that focus on an individual's *thoughts about* knowledge such as their own certainty about the information, what they know about the information source, or if they plan to test the validity of the information, would be considered epistemic metacognition. Delineating epistemic cognition and metacognition prevents conflation of these two constructs, and allows greater insight into students' epistemologies.

2.7.3. The Toulmin Argument Pattern

Arguments were analyzed to characterize the epistemic practices participants use to construct knowledge. The Toulmin Argument Pattern (TAP) guided my analysis of participant arguments. TAP describes the structure of an argument as containing evidence leading to a claim, whose connection is explained through a warrant (Toulmin, 2003) (Figure 2.2). Following the AIR model for epistemic cognition (Chinn et al., 2014), an individual evaluating or constructing an argument with an epistemic aim would be using epistemic process(es). Observing how an individual determines the reliability of evidence, and explains how the evidence relates to claims through the use of warrants reveals the kinds epistemic processes the individual is using. We can utilize these processes by working backwards to infer the ideals individuals are trying to achieve. These interpretations were triangulated with interviews in which we asked students about the epistemic ideals they are trying to achieve. Finally, we incorporated the aims and values portion of the AIR model by asking students why the arguments are important to them, and triangulated that information with the ideals and processes data.

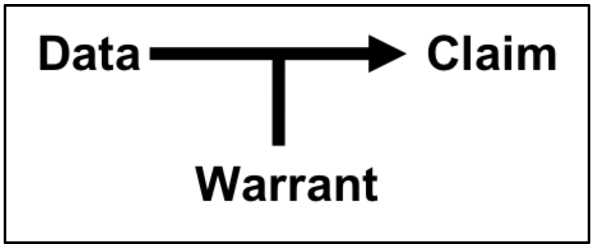

Figure 2.2. A simplified diagram of the Toulmin Argument Pattern (TAP) (Toulmin, 2003).

3. PILOT STUDY

This chapter describes a pilot study to establish and test methods described in Chapter 4. Arguments presented in participants' research papers were analyzed to uncover the epistemic practices they were using to build the argument. I begin the chapter with an overview of the goals of the pilot study, followed by a description of the framework I used to guide my analysis of student arguments. Next, I briefly describe my data collection and analysis methods. I present my results in the form of a model describing the variations in how students consume and produce knowledge, as well as a discussion of how participants' need for closure influenced their argument. I end this chapter with a discussion of how results from this pilot were brought forward into the main study.

There were three primary goals for the pilot study:

1. Establish a method to identify and code students' written arguments
2. Identify sources students use to provide data/evidence for claims
3. Develop an interview protocol to explore the epistemic practices students use to build arguments

3.1. Argument Framework

The essential components of an argument are the claim and the data that support this claim. The claim begins as an assertion, that is, a statement that one expects someone else to take seriously. If the assertion is challenged, it becomes a claim because its validity has been questioned. One must defend this claim (C) by producing facts or data (D) to support the assertion, forming the foundation of the Toulmin Argument Pattern (TAP) (Toulmin, 2003). However, if an assertion is made without data, it is not an argument since the individual making the assertion sees no need to show the validity of the statement through the presentation of data. For example, a simple argument may consist of a claim, “You cannot park in this space” and the datum, “When I parked there I got a ticket.” If the datum is omitted, the argument becomes a statement, “You cannot park in this space.” Following this logic, I used claims and data as a first step to identifying arguments in student papers. For each student paper, two coders highlighted blocks of the paper and identified arguments based on a claim, and data related to the identified claim. To determine if data were related to a claim, we referred to a third component of the TAP, the warrant (W).

A warrant is produced in response to the question “How did you get to the claim from these data?” In our parking example, our warrant would be “Penalties are issued when you are parked where you are not supposed to.” Thus, our more complicated argument would now read “When I parked here I got a ticket (D) *so* you cannot park in this space (C) *since* penalties are issued when you are parked where you are not supposed to (W).” Note that this warrant can connect other statements to this argument as data. For example, the data “Bob’s car was towed when he parked in this space” and “Janet’s car had a parking boot put on it when she parked here” could be connected to the claim “You cannot park here” through this warrant. Support for the warrant comes in the form of backing.

Backing for a warrant presents information that supports a warrant. For example, support for the warrant “Penalties are issued when you are parked where you are not supposed to" may be documents that show the penalties for illegal parking, or legislation that establishes the illegality of parking somewhere. This is where Toulmin (2003) says the argument becomes discipline specific. The backing of documents or legislation showing the penalties uses legal evidence, whereas "here are all the checks I wrote to the DOT for parking tickets" may be a backing based on financial evidence.

3.2. Data Collection

To pursue the three goals outlined above, I sought a group of students similar to the target population, namely those enrolled in an introductory biology course. I found a course at a large, southeastern land grant institution that required students to write a literature review on a controversial topic for the laboratory portion of an introductory biology course. Topics were chose by students, approved by the graduate teaching assistants in each section, and students were given approximately two weeks to research and write the papers. I collected students' literature reviews, identifying and analyzing the arguments each student made in their essays.

3.3. Research Paper Coding

Coding of student papers was done in two phases. The first phase identified the parts of the paper that represent arguments. This was done through identification of the components of argument according to the TAP (Toulmin, 2003). The second phase identified the epistemic practices the student used to present a claim, support the claim through the use of data, and explain the connection between data and claim through a warrant and backing.

The process for coding the research papers grew out of Toulmin's (2003) distinction between discipline specific and non-discipline specific aspects of an argument. Since backings for warrants were discipline specific, backing became one focus of our analysis. Connected to the backing were data that were presented for the argument. The kinds of data and ways that the data are used to support a claim are also influenced by the epistemic cognition that individuals use. Thus, the initial identification of claim, data, warrant and backing was integral to analysis of arguments within participants' research papers.

For this study, a claim was identified in concert with its supporting data, since an assertion made without evidence is not considered a part of an argument (Toulmin, 2003). Data were considered to be supporting a claim if an explicit warrant is stated, or an implicit warrant could be constructed by the researcher to connect the data to the claim. To ensure theoretical validity, implicit warrants were only constructed for data presented in the same paragraph or logical structure of the claim unless the data were explicitly connected to a distal claim by the author. To ensure interpretive validity, claims, data and warrants of arguments were discussed between two coders, who reconciled any disagreements before moving onto the second phase of coding.

Before coding, the researchers wrote memos explaining the kinds of sources they used for data, and how they warranted those data to make scientific claims. This helped to explain how the coders perceived the arguments by making explicit the coders' positionality as scientists. The memos detailed the perspective the coders used to interpret the arguments presented by the student papers. Since the intent of the authors' argument is unknown, the main study employed interviews to probe the ontological epistemic practices students use to build arguments.

3.4. Results from Research Paper Pilot Study

Preliminary analysis of student research papers suggested that students both consume and produce knowledge in their research papers. This production and consumption of knowledge can occur using both scientific and non-scientific epistemologies. In this pilot study, I viewed the biology classroom as a community of practice, where students and instructors share common practices of biology (Lave & Wenger, 1991). These practices include ways of consuming and producing scientific knowledge. This community of practice is connected to epistemology in that students that consume and produce knowledge using a scientific epistemology are more central in the scientific community of practice. An illustration of the Epistemic Consumption and Production of knowledge (ECaP) model can be seen in Figure 4.1.

Consumption of knowledge refers to how an individual finds information and determines that it is fit to be considered knowledge. The ways in which an individual determines if something counts as knowledge is dependent on epistemic cognition, and is therefore context and discipline specific (Chinn, et al., 2014; Louca, et al., 2004). Scientists expect that any claims are supported by either observational or experimental evidence. These data are accessible by observation and therefore multiple observers should be able to come to consensus relatively easily (Norman G Lederman, 2007). Therefore, individuals who use a scientific epistemology will evaluate information based on experimental evidence, observational evidence, and expertise of authors. Such information will generally be drawn from original research or scientific reviews published in peer-reviewed journals. In the excerpt below, participant S037 presents data after describing the original research experiment that produced the data. These data were drawn from a peer-reviewed article from the medical journal titled *The Lancet*.

> *In a study which examined the effect of controlled-release melatonin on the quality of sleep in 12 subjects, all 76 years of age, who reported trouble sleeping, it was found that the maximum level of secretion of 6-sulphatoxymelatonin (aMT6), a melatonin metabolite that helps to control circadian rhythms, was below normal. For three weeks, subjects were given 2 milligrams of controlled-release melatonin each night; for another three weeks, they were given a placebo. Sleep quality was observed using an actimetry sensor worn on the wrist of each subject which monitored movement. The results of the experiment showed that sleep efficiency was 12% greater after melatonin treatment than after the placebo (p<0.1%), and wake time after the onset of sleep was also shorter by 24 minutes when treated with melatonin (p<0.1%) (Garfinkel et al., 1995). –S037*

On the other hand, individuals who use a non-scientific epistemology may not understand the difference between inference and empirical evidence (Norman G Lederman, 2007). Therefore, they may evaluate information based on alignment with political or religious views. These individuals may draw information from sources that explicitly state their affiliation or bias toward a political or religious viewpoint. In the excerpt below, participant S010 draws information from the “Human Coalition,” a group that works to limit abortion through outreach to church communities.

> *A big saying has emerged over the years, “life starts at conception.” This phrase is said by some pro-life believers to express that they believe the moment the sperm fertilizes the egg is when life beings. According to Fisher, the moment that an egg is fertilized, a human begins to grow. From that very moment, the embryo utilizes all the DNA from the egg and sperm and beings building itself into a human (Fisher, 2016).*
>
> *Fisher, Brian. "The Points of Proof for Life: Scientific." Human Coalition. 12 May 2016. Web. 28 Mar. 2017. –S010*

Production of knowledge refers to how an individual constructs the conclusions that are drawn in the research paper. An individual producing knowledge with a scientific epistemology will make claims supported by citation of experimental or observational evidence. The individual will extrapolate their conclusions through a synthesis of the data presented. In the excerpt below, participant S037 makes the claim that melatonin supplements help those with sleep-related disorders, supporting the claim with experimental data and an explanation of how the experimental data were produced.

> *In another study performed in Finland, 15 children with sleep disorders associated with Asperger syndrome were treated with 3 milligrams of melatonin a day for 14 days and assessed with actigraphy for sleep quality and level of tiredness. The results of the study determined that sleep latency in the subjects had been reduced from 40.02 ± 24.09 minutes to 21.82 ± 9.64 minutes (p=0.2%), while total duration*

of sleep remained unaffected (Paavonen et al., 2004). Overall, melatonin supplements appear to have a positive effect on those who suffer from sleep-related disorders. –S037

Individuals who consume knowledge in ways that align with science epistemology may yet produce knowledge in ways that do not align with science practice. For example, participant S071 cites only peer-reviewed literature, but bases their conclusions on perceptions about the level of patient understanding.

Another reason for not seeking consent is the patient's level of understanding of what is being discussed. Even though patients know the circumstances of obtaining an HIV infection, they would probably misunderstand if the infection is presented to them. Many questions were asked about the community staff containing information about an infected patient, for this information could benefit the patient, doctor, or nursing staff. The staff involved with infected patients should only be informed about their infection. The patient should be told about what members of the staff associated with him/her have this confidential information. –S071

Additionally, an individual using a non-scientific epistemology may make claims without citing experimental or observational evidence. As such, the individual may report information with little synthesis, presenting little or no explanation of how the information was produced. In the excerpt below, participant S030 presents multiple statements about hormone function.

Hormones have a huge effect on the body during pregnancy. After pregnancy, your hormones will go back to being normal overtime [sic]. The estrogen and *progesterone levels will decrease and the body will go back to the normal levels over time. –S030*

While S030 presents these statements as fact, they do not support these statements with experimental or observational evidence, as is customary in science. As such, we consider this statement as non-scientific knowledge production.

Analysis of student research papers revealed that they consumed and produced knowledge in ways that both aligned and did not align with the scientific community of practice. From this analysis, I constructed the Epistemic Consumption and Production

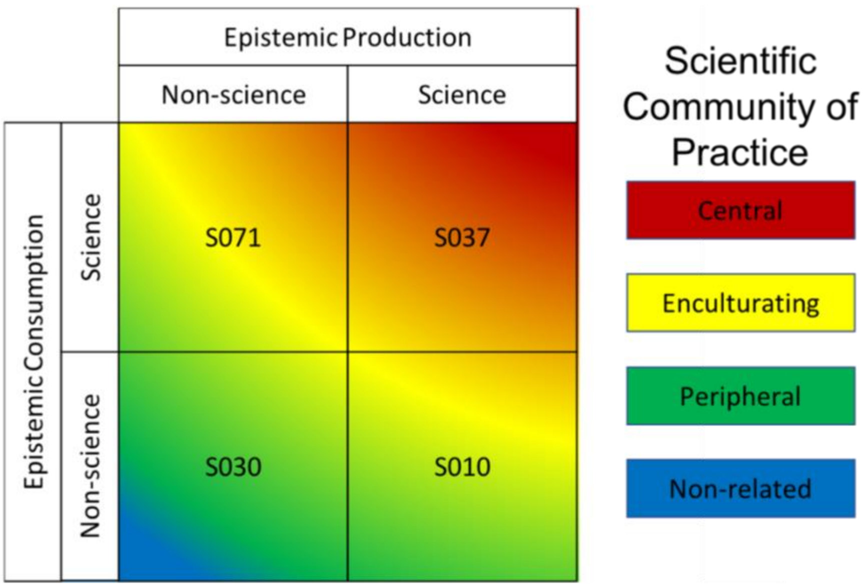

Figure 3.1. Model for student epistemic consumption and production (ECaP). Proximity to the center of the community of practice is represented by a color gradient from red to blue, with red being most central and blue being outside the community.

(ECaP) model to represent how each participant consumed or produced knowledge. I considered participants that both consumed and produced knowledge in scientific ways to be most central in the scientific community of practice. I considered participants who used either non-scientific knowledge consumption or production practices to be non-central. Therefore participants S010 and S071 are placed in quadrants that are colored yellow and green. Finally, I considered participants who both consumed and produced knowledge in non-scientific ways to be most distant from center. For this reason, participant S030 is placed in the quadrant colored green and blue. The ways in which students consume and produce knowledge can affect how they build arguments in biology classrooms. Therefore, this model will be helpful for assessing how participants in my main study consume and produce knowledge to ensure that there is maximal variation in my sample.

3.5. Need For Closure

In addition to the epistemic consumption and production of knowledge, we also found that participants' need for closure influenced their arguments. Need for closure is part of the Lay Epistemic Theory, which describes how an individual comes to know something through their cognitive capability and epistemic motivations (Kruglanski, 1990). The epistemic motivations are driven by an individual's need to either pursue or avoid specific and non-specific closure. An individual seeking specific closure seeks a particular solution, while an individual seeking non-specific closure would accept any solution (Kruglanski, 1990). Individuals with different needs for closure would approach arguments in different ways. For example, an individual avoiding specific closure aims to avoid coming to a particular conclusion, and may therefore present that particular conclusion in an unreasonable light. Participant S010 from the research paper pilot study avoids the specific closure of legal abortion by presenting the viewpoint through the lens of "Pros of Abortion." Participant S010 provides evidence for why some forms of abortion should be legal, including cases in which the mother's life is in danger. However, Participant S010 disregards the evidence in favor of the following statement, which provides an unreasonable stance that allows Participant S010 to avoid specific closure.

> *Regardless of rape, genetic predispositions, and other reasons for abortion, some people believe that women should have full range to do what the wish with their bodies. These people advocate "pro-choice," where the woman can choose whether or not to terminate her child for whatever reason. They believe if the woman doesn't want the child, she should be able to abort it no questions asked. – Participant S010*

Conversely, an individual avoiding non-specific closure would avoid making any strong arguments. Participant F095 provides evidence about a theory of mitochondrial evolution, but immediately follows with a statement about the lack of reliable evidence.

> *How and why mitochondria actually evolved is hard to be certain of as it took billions of years to accomplish. Unfortunately, there are also no fossil records for individual cells, especially those as small as the mitochondria's ancestor. There is no way to possibly track the course of its physiological changes over vast amounts of time. –Participant F095*

These two participants are motivated by avoidance of specific and non-specific closure, respectively, and as a result present two different argument strategies. This suggests that choosing participants with different needs for closure can enhance the variation of epistemic practices the study population uses to construct arguments. Webster and Kruglanski (1994) developed and validated an instrument to measure an individual's need for closure, and found that the items loaded into five factors. However, I will only be using items that test two of those factors, discomfort with ambiguity and closed-mindedness, because they have been described as epistemic vices (Chinn, *et al.*, 2011), a component of the theoretical framework used in this study, the AIR model for Epistemic Cognition (Chinn, *et al.*, 2014).

3.6. Discussion and Conclusion

The purpose of this pilot study was to establish a method to identify and code students' written arguments, identify sources students use to provide data/evidence for claims, and develop a method for participant selection. TAP was an effective way to identify participants' arguments, but this pilot study revealed that participants seldom warranted their claims. Since the warrant explains the connection between data and claim, the warrant often reveals the epistemic practice(s) the arguer is applying to the argument. The absence of a warrant therefore makes it difficult to determine the applied epistemic practice(s) without extensive interpretation by the coder. As such, I plan to identify the parts of an argument before interviewing participants so if warrants are missing, they can be identified during the interview.

Analyzing student arguments also gave insight into the kinds of sources students use to support their claims. Identifying these sources helped to determine whether students' epistemic consumption aligned with the expectations of science. In the main study, I continued to analyze participant sources of information. I also integrated questions about why participants sought specific sources of information to support their claims in my interview protocol. This protocol was further refined in pilot interviews described in the methods section (Chapter 4).

Finally, the pilot study describes a variety of ways participants consumed and produced knowledge through the ECaP model, and revealed that participants' need for closure affected their argument construction. Since participants placement on the ECaP model and variations in need for closure was correlated with different kinds of arguments, these two criteria will be used as part of the selection criteria for participants.

4. METHODS

4.1. Institutional Review Board Approval

This work was approved by the University of Wisconsin Institutional Review Board (IRB) under the approval number IRB-2018-1023, and the Clemson IRB under the approval number IRB-2016-244. All recruitment emails, scripts, interview protocols, and statements of consent were approved through these IRBs, and are included in Appendix A.

4.2. Research Objectives

The end goal of this project is to construct a framework that can be used to categorize the kinds of epistemic practices students use to build arguments in biology. To this end, this project will address the following research question:

> What are the qualitatively different ways that students use epistemic practices to build an argument in a biology course-based undergraduate research experience (CURE)?

In this study, I use a phenomenographic approach (Åkerlind, 2012; Green & Bowden, 2005; Marton & Booth, 1997). Marton and Booth (1997, p.111) describe phenomenography in this way:

> At the root of phenomenography lies an interest in describing the phenomena in the world as others see them and in revealing and describing the variation therein, especially in an educational context…

Central to phenomenography are the ways in which something is experienced. To experience a phenomenon, one must both discern its *structure* from the surrounding context, and assign *meaning* to that structure (Marton & Booth, 1997). In the context of the biology CURE assignment, we ask the participants to discern the structure of an argument and assign meaning to it. These two aspects of the experience are intertwined: the meaning of the argument is derived from how the structure of the argument is organized, while the structure of the argument is driven by the meaning that the participant intends. Student epistemic practices are embedded in the decisions to include, exclude or combine aspects of an argument structure with the intent of clarifying the meaning of the argument. The different ways that a phenomenon is experienced is reported as “categories of description” (Marton & Booth, 1997) that describe the different meanings and structures that emerge as a part of that experience. The categories of experience are then situated in an “outcome space” that represents the relationships between the categories (Åkerlind, 2012).

The focus on revealing the variation in how others experience a phenomenon in the world also makes phenomenography a good fit for this study. An important aspect of this study is the focus on the student perspective on building arguments in a biology CURE. Using phenomenography I organized students’ epistemic practices into categories of descriptions (Marton, 1986). Green and Bowden (2005) suggest that phenomenography should be guided by both the principles of phenomenography and the practical issue that the researcher is attempting to solve. As such, while constructing categories of descriptions and outcome spaces, we should also think beyond to practical issues such as improving instruction to help students think more like a biologist.

There are several methods that I could use to investigate the ways that students build arguments in a biology CURE such as grounded theory, phenomenology and phenomenography. Grounded theory would be useful in constructing a theory about how students build arguments, however in the class that I studied, students received feedback every time they wrote a research proposal. In grounded theory, participants are purposely selected to inform developing areas of the model over time (Charmaz, 2014). However, in my study context, students' conceptions of building an argument likely changed over time due to instructor feedback. Since I used the research proposals as sources of data, this complicated the execution of the constant comparative method. Therefore, grounded theory was not an appropriate method for this project.

On the other hand, phenomenology is an attractive methodological approach because it focuses on the essence of how participants experience a phenomenon (Creswell, 2012), in this case, building an argument. However, phenomenology results in the description of one construct that represents the essence of how students experience a phenomenon. When thinking about the practical issue of teaching students about thinking like a biologist, understanding the diversity of how students experience building an argument (as in a phenomenography) provides a broader perspective of the issues that need to be addressed while planning lessons.

4.3. Study Context

Studying epistemic cognition is difficult because for an individual, epistemic cognition is often tacit (B K Hofer, 2004). Furthermore, the contextual nature of epistemic cognition (Chinn et al., 2014; Elby & Hammer, 2001; Elby, *et al.*, 2016) necessitates that I examine epistemic cognition in an authentic environment. Therefore, I investigated students' epistemic practices while they participated in a core scientific practice: constructing an argument.

Participants were recruited from an introductory undergraduate biology course in the Biology Core Curriculum (Biocore) at the University of Wisconsin-Madison. Biocore is a four-semester honors biology curriculum that students typically begin sophomore year, but some freshmen are allowed entry pending advanced placement credit that fulfills prerequisite course requirements. Students must apply and be accepted to Biocore to enroll in Biocore courses. The first course in the Biocore sequence was chosen because it satisfies the CURE guidelines established by Auchincloss *et al*. (2014; see "Classroom/Course selection" below), and because it provides students with opportunities to engage in argumentation through course activities and assignments.

During the semester, students complete three projects following a consistent sequence of activities: generating a hypothesis, proposing a research plan for critique by classmates, collecting, analyzing and interpreting pilot data, and writing a revised research proposal based on the results and interpretations. Previous students have used these research proposals as a starting point for independent study with a faculty mentor. The structure of the course integrates all five activities that help define a learning experience as a CURE: *use of scientific practices, discovery, broadly relevant or important work, collaboration,* and *iteration* (Auchincloss et al., 2014; Brownell & Kloser, 2015). Students are encouraged to *use scientific practices* such as making observations to formulate testable hypotheses, designing and conducting investigations, analyzing and interpreting data, constructing new knowledge, and communicating their findings to a broad audience. Student projects involve work in a faculty and student maintained Biocore prairie restoration project, which makes the work *relevant* to the community beyond the classroom. Projects are proposed by the students, as such, the results are not known to instructors or students, providing opportunities for *discovery*. Students must propose their research in teams to the rest of the class for critique, resulting in multiple levels of *collaboration*. Finally, the requirement of students to revise their research proposals, perform experiments outside of class time, and complete multiple projects, provides students with a structured environment for *iteration*.

Furthermore, this course structure offers opportunities for students to engage in the construction and evaluation of arguments. Proposing the research plan to the class requires that the authoring students present an argument supporting their chosen research plan, while their classmates must evaluate that argument to provide critique. Students must construct a revised argument when writing their final research proposal, and must evaluate the arguments of other students when providing feedback during peer review. This provides an environment where students and instructors co-construct a catalog of epistemic practices that are acceptable in biology through discussion, peer review, instructor guidance, grading, and feedback. The integration of all five CURE activities and the opportunity for students to engage in argument make this course an excellent setting for my study.

4.4. Research Quality Considerations

The Quality Framework (Q3) developed in engineering education (Sochacka, Walther, & Pawley, 2018; Walther, Sochacka, & Kellam, 2013) provided the language to describe and guide my thinking on key research quality issues throughout data collection and analysis. Q3 conceptualizes interpretive research quality issues as six constructs: Theoretical, Procedural, Communicative, Pragmatic, and Ethical Validity, and Process Reliability. Theoretical Validation addresses the ways which ensure the theory we generate from our analysis is representative of the social reality. Procedural Validation addresses how the research design ensures that knowledge built from the project aligns with the social reality being studied. Communicative Validation concerns the ways in which data and analyses are effectively communicated between members of the research group, discipline, and beyond. Pragmatic Validation addresses the compatibility between the theoretical framework(s) and the social reality under investigation. Ethical Validity considers the underlying human elements that govern the influences between researchers and participants. Finally, Process Reliability ensures that the processes used in the project are dependable and consistent. Each of these constructs is embedded throughout the research process, rather than focusing only on the quality of research outcomes (Walther et al., 2013). A summary of how the research quality framework was embedded throughout the research project is presented at the end of this chapter in Table 2.

4.5. Data Collection

4.5.1. Participant Selection

The focus on variation of experiences in phenomenography necessitates the purposive sampling of participants for maximum potential variability in experiencing a phenomenon (Bowden & Green, 2005). Seeking maximum variability in the sample mediates threats to theoretical validity by ensuring the dataset includes an approximation of the variability seen in social reality. For this study, the phenomenon in question is using epistemic practices to build a scientific argument in a class, so criteria used to select participants were focused on epistemic processes that affect argument construction, arguments made in past assignments, and the course section in which argument construction occurred. I used the Epistemic Consumption and Production (ECaP) model (described in Chapter 3) in concert with the comfort with ambiguity and closed-mindedness factors from the Lay Epistemic Theory (Webster & Kruglanski, 1994) to identify possible variations between participants' epistemic practices. Inclusion of the ECaP model and Lay Epistemic Theory in participant selection address concerns about pragmatic validation by ensuring that meaningful epistemic insights can be drawn from this sample.

The ECaP model was constructed by analyzing different arguments written by introductory biology students (Chapter 3). In this study, we found that students both consumed and produced knowledge in scientific and non-scientific ways. Since students varied in their consumption and production of knowledge in writing arguments, selecting for participants that displaying a variety of knowledge consumption and production strategies would yield a greater diversity of epistemic practices. The way that students consume and produce knowledge suggest that they hold certain epistemic beliefs. For example, a student that consumes knowledge by believing anything written by an expert may believe that experts are the sources of knowledge. This may affect how they present evidence in an argument, such as citing expert conclusions without explanation. Conversely, a student that produces knowledge through reaching consensus may believe that knowledge should be justified through peer-review. This student may consequently support arguments by presenting experimental evidence from peer-reviewed publications. Thus, selecting participants from different quadrants of the ECaP model would increase the variability of epistemic practice use while constructing arguments.

I administered the Need for Closure instrument including only the items related to comfort with ambiguity and closed-mindedness to ascertain students' need for closure. I chose to include these two constructs because they have been previously identified as epistemic virtues and vices related to epistemic cognition (C. Chinn et al., 2014). Additionally, I analyzed students' first proposal to determine how students fit into the ECaP model, choosing representative students from all four quadrants of the model. Finally, student epistemic practices can be influenced by instruction. I recruited students from all sections of the course to ensure that students taught by each graduate teaching assistant was represented in the sample. Thus, participant variability was ensured along three dimensions: course section, need for closure, and epistemic consumption and production.

4.5.2. Description of Participants

Thirty-two students submitted their first research proposal assignment to be analyzed, and of those students, twenty-one participants agreed to submit their second research proposal assignment and be interviewed. I gathered demographic information from these twenty-one participants during the interview, asking about their year in college, major, gender identity, racial identity, ethnic identity, and if they were a first generation college student. To preserve participant anonymity, I report participant demographics here in aggregate. Most of the participants were sophomores (*n*=18), while the rest were freshmen (*n*=3). Almost all participants were majoring in a life-sciences related field (biology, biochemistry, genetics, molecular biology, or neurobiology), while one student was double-majoring in psychology/biology, and another was majoring in communication sciences and disorders. About three-quarters of the participants identified as female (*n*=16), and the rest of the participants identified as male (*n*=5). One student specified herself as cis-female. More than half of the participants identified as White or Caucasian (*n*=12), followed by Asian (*n*=5), Hispanic (*n*=1), Indian/Bengali (*n*=1), Jewish (*n*=1), and Thai/Middle Eastern (*n*=1). Two participants reported that one parent finished college, and one participant identified as a first-generation college student, while the rest of the participants reported that both parents were college graduates.

The criteria I used to identify the variability in the sample population are presented in Table 1. I checked the internal consistency of the items for the NFC constructs Discomfort with Ambiguity and Closed-mindedness through Cronbach's alpha, obtaining an acceptable reliability coefficient for Discomfort with Ambiguity (α=0.74). The reliability coefficient for the Closed-mindedness factor was not acceptable (α=0.45), and therefore was excluded from participant selection. Participants rated their Discomfort with Ambiguity from 1 (comfortable with ambiguity) to 7 (uncomfortable with ambiguity), and their ratings ranged from 3.13 to 6.38. All quadrants of the ECaP model were represented in the sample population. *n*=8 participants fell into quadrant 1, *n*=5 participants in quadrant 2, *n*=7 were in quadrant 3, and *n*=1 participant was in quadrant 4. All four course sections were represented in the sample population. Section numbers were replaced with letters to protect the identity of participants. The most participants were from section A (*n*=8), followed by section B (*n*=5), while C (*n*=4) and D (*n*=4) had the same number of participants.

Table 1. Variability of Sample Population

Pseudonym	Disc. w/Ambig.	ECaP Quadrant	section
Alyssa	4.00	3	D
Bill	3.88	3	B
Bonnie Clyde	6.38	3	A
Cat	4.50	1	A
Claire	4.75	3	B
Connie	5.00	2	A
Dan	3.25	1	A
Ella	5.50	1	B
Flo	5.75	1	A
Hannah	5.00	1	C
Jane	4.25	2	A
Janice	4.00	4	D
Jennifer	5.00	2	D
Joe	4.88	1	C
Karen	6.38	2	C
KBrown	5.50	3	C
Mary	4.75	1	A
Megan	5.38	1	B
Nick	6.38	3	B
Sam	5.88	2	D
Surprise	3.13	3	A

4.5.3. Proposal Argument Identification

In light of the difficulty interpreting students' epistemic practices solely from their research papers in the pilot study, I included student interviews as part of the data collection plan (See Appendix A for interview protocol). Including interviews in data collection addresses ethical, procedural, and communicative validation since the interviews give researchers an opportunity to probe students on possible epistemic practices identified from students' written work, and to represent student voices via quotes from the transcript. However, to interview students about their argument, we must first have some understanding of the argument they present. Therefore, student research proposals were analyzed using the Toulmin Argument Pattern (TAP) (Toulmin, 2003) prior to collection of interview data. This initial analysis allowed me to address pragmatic validity by asking interview questions that specifically probed the reliable processes, or epistemic practices, students use to build their argument. Analysis of the student research proposals before interviews was a necessary departure from phenomenography, which typically calls for analysis only after all data have been collected (Akerlind, 2005). This is done to ensure that interviews are conducted consistently, without undue influence from analysis of data already collected. Since initial analysis was needed for me to properly probe participants' epistemic cognition, and no more analysis was done between the start and conclusion of interviews, I deemed this practice an acceptable limitation to my study.

As a part of the research proposal, the instructors require that students explain the biological rationale behind their experimental plan. This biological rationale explains how students' experimental plan addresses the research question. From the perspective of TAP

(Toulmin, 2003), the *claim* would be that the experimental plan adequately addresses the research question, the *data* would be information gathered from sources (experimental, literature search, etc.), and the *warrant* would be the biological rationale. This fit between TAP, research proposal and biological rationale demonstrates theoretical validity as TAP is sufficient to describe and identify student arguments within the research proposal (Walther *et al*., 2013). Similar to the methods described in the research paper pilot study, student research proposals were coded by examining the kinds of information sources students used to support their biological rationale, and how students used their biological rationale to support the claim that their experimental plan addresses their research question. Students completed three research proposals during the semester, and I analyzed the second proposal each participant wrote. I chose the second proposal because students were more familiar with the process of constructing the proposal, and the first proposal was used as part of the participant selection process. I did not choose the third proposal because was presented as a group, making it difficult to parse individual epistemic practices from the group project.

4.5.4. Interview

Interview data is traditionally used as the primary data for phenomenography (Bowden & Green, 2005). In phenomenographic interviews, participants are asked about an initial scenario in order to keep data collection consistent between participants and interviews (Bowden, 2000; Marton, 1986), contributing to process reliability (Walther et al., 2013). Consistency between interviews is important for phenomenography because the variation in the ways participants experience a phenomenon should come from the participants themselves, rather than the way data were collected (Marton, 1986). For these reasons, pilot interviews are often conducted to ensure that the opening scenario consistently elicits proper data from the participants (Bowden & Green, 2005). Furthermore, these pilot interviews give the interviewer an opportunity to practice probing consistently and effectively with follow-up questions, contributing to both process reliability and procedural validation (Walther et al., 2013).

I piloted the interview protocol with four undergraduate research experience (URE) students at Clemson University. All interview pilot participants chose their own pseudonyms. I chose to interview URE students because their experience is similar to the experiences of students in the Biocore program. Students in both programs are asked to justify their research plan and conclusions in a poster format. For each participant, I analyzed their poster by identifying an overall claim, data, and sub-arguments that supported claims throughout the poster. I then wrote follow-up questions to identify warrants, backing, and epistemic practices students used to build their arguments. As the initial scenario, I asked the first interviewee, Elizabeth, "Could you go through your poster for me?" with the expectation that she would state her claim and data for her overarching argument. However, it was difficult to identify her overarching claim as I listened to her presentation. For the following participants, I changed the opening scenario question to "What are you trying to get across with this poster?" This question consistently elicited the overarching claim I identified when analyzing participants' poster prior to the interview for the rest of the three pilots. Additionally, I found myself tempted to ask students about practices that I would have used, such as “Why didn’t you do this…?” Writing the positionality memo described in the research paper pilot alerted me to my own practices and helped me realize that I would have been adding content to the interview independent of the participant. Preventing myself from adding content to the interview helps with communicative validity, ensuring that any practice mentioned by the participant comes from their experience of the phenomenon, not from content provided by the researcher (Bowden & Green, 2005).

Phenomenographers suggest completing 20-30 interviews to ensure that there is enough variation in the ways of experiencing a phenomenon, without generating an overwhelming amount of data (Bowden, 2000). Thus, I aimed to invite 30 participants for interviews, understanding that some potential participants may choose not to be interviewed. In the end, twenty-one participants agreed to be interviewed. Interviews were done in a short timeframe to limit time as a factor in changing perceptions. In my case, interviews were completed within one month after collection of research proposals, as the next research proposal assignment was due at that time.

I collected all participant proposals two weeks before the interview period to review content and plan follow-up questions for the interview. To aid in initial analysis, I constructed a poster analysis template (Appendix D). This template document provided space for me to list the hypotheses proposed by each participant and any data/warrants for these hypotheses. The document also provided space for me to write any follow-up questions for the anchoring questions asked during the interview.

The interview was anchored around 4 main questions, listed below.

Main Questions:

1. What is the importance of this project to you?
2. How did you come up with your hypothesis? Where did you get information for your hypothesis?
3. What conclusions did you draw? How did you know they were correct?
4. Are there parts of your project that you would consider knowledge? What makes those components knowledge?

Question 1 was drawn from the AIR model's Aims and values component (Chinn, 2014) to establish the participant's aim for the project. Upon reviewing all participants' proposals, it became apparent that they followed a consistent template. Each participant included introduction, hypothesis, expected/alternative results, methods and implications sections. Question 2 was designed to probe arguments made in the hypothesis, expected/alternative results and methods sections. While I expected students to use the biological rationale to warrant their hypotheses, I did not explicitly ask about the biological rationale unless the student mentioned it. This was done to ensure that the discussion of the biological rationale emerged from the student rather than my questioning. Question 3 emerged from the first interview with a participant. It was originally intended to probe participant arguments in the implications section, but the first participant felt more comfortable discussing the actual results of the experiment instead of the implications. I continued to ask students to explain their implications section, but all participants transitioned to talking about their results at this point in the interview. The final question was designed to probe participants' epistemic ideals (Chinn, 2014). This question was asked at the end of the interview so participants could refer back to our previous discussions to construct their answers. Participants seemed to find this the most difficult to answer, and indeed many did refer back to previous discussions to aid in their thinking for this question.

Following each interview, I wrote a memo describing the interview experience with the participants. In writing these memos, I focused on the epistemic practices that were apparent to me during the interview. I also commented on my perception of interview quality and noted any improvements that could be made for the next interview. Finally, I included the measures for participant selection, namely the participant's Discomfort with Ambiguity rating and ECaP quadrant.

4.6. Data Analysis

Participant assignments and interviews were analyzed through a combination of initial coding (Charmaz, 2014) and process coding (Saldaña, 2016). I used the AIR model for epistemic cognition to guide my analysis, since a framework for epistemic cognition specific to the context of biological thinking has not been established. Initial coding is useful because it allows meaning to emerge from the data without being bound to pre-existing categories, while the AIR model helps to focus the analysis on epistemic concepts. The gerund use in process coding fits this project because it identifies actions that participants take. This will help to identify the epistemic practices participants use during the construction and evaluation of arguments. The coded actions of participants will be organized into processes. I also wrote memos detailing the context, analysis decisions and meaning making that I did during the data analysis, as described by Charmaz (2014). During this process, I also looked for feedback from my committee members, fellow graduate students who have expertise in qualitative methods, Biocore instructors and participants to ensure that construction of the outcome space is as close to reality as possible. In the following sections, I will outline some of the details of my analysis by using some of the guiding questions outlined in Akerlind *et al.* (2005).

Analysis began once interviews were complete

Analysis of the interviews began after all data were collected. This was done to prevent any analysis from influencing later interviews (Akerlind et al., 2005). As stated above, consistency between interviews is of paramount importance since variations between the data should come from the participant rather than from differences in data collection (Marton, 1986). Once analysis began, participants were evaluated in the context of all other participants, as in a phenomenography, an understanding of the categories of description can only be achieved through analysis of the participants as a whole, not through individual participants (Bowden, 2000).

Data was analyzed in the context of the transcript

Akerlind (2005) suggests analyzing the data either in the context of the transcript, or by identifying excerpts and pooling them to remove them from the context of the transcript. While pooling the excerpts helps the researcher to see the data in the context of all participants as one (Akerlind, 2005), my project required me to keep the excerpts in the context of the transcript because of the context dependency of epistemic cognition (C. Chinn et al., 2014; Louca et al., 2004). Furthermore, the epistemic practices are identified in light of the underlying context of the claims, data and warrants presented within each participant's research proposal. Thus, each excerpt needed to be associated with the participant that generated it.

4.6.1. Team Analysis

The analysis was done as a team to collectively build the categories of description and outcome space as the negotiation between multiple viewpoints of researchers coming to consensus helps to temper the biases of individual analysts. Indeed, Bowden (2005) suggests that team analysis ensures that individual researchers bracket their own perceptions of the phenomenon, and that the evidence used comes exclusively from the transcripts. Bracketing and drawing evidence only from transcripts address procedural, communicative and pragmatic validation (Walther *et al.*, 2013) as they ensure the conclusions constructed from the data reflect the participant's experience rather than the experience of the researcher. While I have adopted bracketing, drawing evidence only from the transcript was not feasible since the participant's research proposal formed the foundational context with which to interpret the participant's experience of building an argument. I do not believe that including the participant's research proposal as a source of evidence will hurt the validity of this study since the proposal is written entirely in the participant's voice, and is, in fact, the product of the participant's experience.

Two analysis team members were trained in phenomenographic methods, and had background knowledge on the AIR model for Epistemic Cognition to ensure that they have both the procedural and theoretical background to make meaningful contributions to the analysis. Understanding of the theoretical framework is essential as team members will be expected to act as devil's advocates, demanding evidence that supports the construction of both the categories of descriptions and outcome space (Bowden, 2005).

4.6.2. Building the Categories of Description and Outcome Space

The products of this phenomenography were built in succession; first the categories of description were constructed, then the categories were organized into an analytic framework to better visualize the relationships between categories, and finally the categories were organized into practice composites. The outcome space describes these composites and the relationships between them. The categories of description depict the participants' experience of the phenomenon. Thus, it is important that the researcher's perspective is kept separate from the construction of these categories. On the other hand, the outcome space requires the researchers' perspective to uncover the connections and relationships between the categories of description. As such, Bowden (2005) suggests that construction of the categories of description and outcome spaces are done separately to prevent contamination of the categories of description with the researchers' perspectives.

To build the categories of description, I followed a modified iterative procedure outlined by Kinnunen and Simon (2012):

1. 1st pass of each transcript: Coders familiarized themselves with the data, and identified epistemic passages
2. 2nd pass: Coders discussed/co-constructed tentative categories

3. 3rd pass: Coders distributed tentative categories, refined categories by reading through transcripts and categorized excerpts

First Pass

During the first pass, three team members read through each transcript to familiarize themselves with the data. This helped the team to have a broad view of the participants as a whole (Kinnunen & Simon, 2012). Additionally, during the first pass, each coder identified passages that they felt were epistemic in nature by finding utterances that related to epistemic aims, ideals or reliable processes, as outlined in the AIR model for epistemic cognition (Chinn et al., 2014). To establish a consistent definition of an "epistemic passage," the three team members initially coded three transcripts. The team went through each identified passage and discussed why the felt the passage was epistemic. As they discussed passages, the team members refined their definition of an epistemic passage. Upon refinement, we agreed upon three criteria to identify epistemic passages:

1. The participant is aiming to understand or generate knowledge
2. The participant is using processes that allow them to collect data, understand content, or generate knowledge
3. The participant is discussing their beliefs about knowledge, knowledge generation, or understanding

After identifying the criteria for an epistemic passage above, we split the remaining transcripts into three groups to perform a round-robin coding method (Faber et al., 2020). The coding was done in three phases: initial, secondary, and check. During the initial phase, each team member coded a clean transcript, highlighting epistemic passages. In the secondary coding phase, the coded transcript was passed to a second team member. This member read through the transcript, checking 1) if passages that were identified as epistemic were indeed epistemic, and 2) if there were parts of the transcript that should have been identified as epistemic and were not. Any coding disagreements between the first and second team member were identified with a code labeled "check." During the final check phase, the initial and secondary coded transcript was passed to a third team member. The third team member resolved disagreements between coders one and two by referring back to the three criteria for epistemic passages outlined above. In cases where the third member could not resolve the disagreements, the passages were discussed among all three coders. In this way, all three team members were exposed to all 21 transcripts during the first pass.

Second Pass

During the second pass, transcripts were distributed among two researchers, and they began constructing tentative categories. At this stage, researchers identified passages that represented participants' epistemic practices. As these practices were identified, they were grouped together without labeling the similarities or differences between categories (Akerlind, 2005). This was done to minimize the cognitive load on researchers as they begin to sort the data. Once categorized, the researchers merged categories as necessary, and collectively described the similarities and differences between the tentative categories.

To construct the initial categories, we chose two transcripts and identified student epistemic practices from passages that were marked as epistemic during the first pass. Once each team member marked initial categories in their respective transcripts, they switched transcripts to apply their initial categories to the other transcript. Following the categorization of these two transcripts, we met to discuss and clarify the categories. These categories served as the initial codebook for coding of the rest of the transcripts. For the remaining transcripts, every week, each team member coded two clean transcripts, then passed the coded transcripts to the other team member with passages marked but without the codes identified. This essentially unitized the transcript for the second coder (Guetzkow, 1950), who applied their own codes to the transcript. Once coding was complete, the two team members met to discuss the transcripts in order to reach inter-rater agreement.

Third Pass

After coding 14 of the 21 interview transcripts, the coding team saw no more emerging categories from the data. At this point, we solidified the descriptions of each category and began refinement. To refine the categories of description, we asked questions like "Does this represent variation that is critical to describing student use of epistemic cognition?" and "Is this category different enough from another to be considered a separate category?" The remaining transcripts were coded with the refined categories.

Once the categories of descriptions were constructed, we began to sketch the analytic framework. Akerlind (2012) describes varying approaches to building the outcome space by either allowing the structure of the outcome space to emerge from directly from the data or from the perspective of the researcher. Construction of the outcome space involves description of the relationships and connections between the categories of description (Marton & Booth, 1997). Since the categories of description emerge from a synthesis of student data (Bowden, 2005), building the outcome space necessitates the researcher's perspective. Therefore, I constructed the outcome space with input from a mixture of student data and my own perspective of biology epistemology.

4.6.3. The Warp and Weft

Sketching the outcome space revealed that the categories of description only described participant experiences at a surface level. While the categories of descriptions adequately described participants' epistemic practices, they did not portray the variation in the complexity of thinking and reasoning described by each participant. To gain deeper insight into participants' use of epistemic practices, I employed the "blenderizing" technique described by Pfirman (2018). I split individual transcripts into units corresponding to the four main interview questions as described above. Analysis then proceeded question by question using the categories established in initial coding. The categories served as a way to assign meaning and importance to each coded excerpt. During this stage of analysis, all parts of the transcripts were coded, not just the passages identified as epistemic, in order to ensure that no meaning was lost from the analysis.

Once coding of all of the deconstructed transcripts was complete, I returned to analysis of each individual participant, moving down their transcripts to look for patterns in how they answered each overarching interview question. In this way, the analysis followed a "warp and weft," where analysis was done across participants for each interview question (warp), then down each individual participant to search for patterns of epistemic practice use (weft). To facilitate the "weft" stage of my analysis, coded passages were pulled from transcripts into individual analysis memos for each participant. The excerpts were then organized under each of the four overarching interview questions. This was done to ensure that participant responses could be analyzed in the context of their entire discussion of each interview question. Once organized, the excerpts under each overarching interview question were used to construct a holistic description of each participant's knowledge process. Attention to the codes assigned to each excerpt was invaluable during this process, as they represented meanings that were connected into a cohesive whole.

While writing the analysis memos, I also printed each of the excerpts included in all the analysis memos to organize them into categories, since the initial categories only

Table 2. Challenges to validity and reliability, and techniques applied to mitigate these challenges to maintain research quality.

		Q3 Quality Framework Constructs						
Challenges	Mitigaion techniques	Theoretical	Procedural	Communicative	Pragmatic	Ethical	Process Reliability	
Is the sample representative of the class population?	Account for criteria that can lead to variability in argument construction	x			x			Making the Data
How do I ensure that my data accurately represent of how students built their arguments?	Use interviews to probe student use of epistemic practices		x	x		x		
How do I ensure that the variability in student use of epistemic practices is not an artifact of data collection?	Ensure that interview protocol, depth of folllow-up probing, and pre-interview analysis is consistent. Perform pilot interviews. Ensure that my follow-up questions do not add content to the interview		x	x			x	
Am I sure that my analysis is not a result of my own idiosyncratic thought process? Does the analysis reflect the social reality under investigation?	Ground analysis in theory: AIR model for epistemic cognition. Code data as a team, making sure they are willing to and feel comfortable questioning each others' analysis. Establish a codebook, check each others' coding to ensure consistent meaning making.		x	x	x		x	Handling the Data
How to I ensure that the analysis stays true to the participant experience and is not unduly influenced by my own?	Ask all coders to memo about their own knowledge beliefs. This way, they are aware of their own beliefs and will be able to bracket more efficiently.		x			x		
Does my analysis adequately represent the diverse ways in which participants used epistemic practices to build arguments?	Go through multiple iterations to ensure no more new insights emerge from the data. Use the "blenderizing" and the "warp and weft" techniques to achieve deeper insights from the data. Contextualize quotes from the data within participant experience by writing analytical memos	x	x			x		

described student epistemic practices at a surface level. Each printed excerpt included the initial category assigned to it during the warp portion of analysis. As I read through each excerpt, I placed excerpts that I felt were similar into groups, without defining the groups (Akerlind, 2005). Once all the excerpts were grouped, I defined the groups according to similarities I saw between excerpts, ensuring that each individual group was unique and could not be merged with another group. Through this process, the excerpts were organized into four categories: gathering information, checking credibility of sources, constructing claims, and supporting experimental claims.

While the four categories of description were better defined than before blenderizing and the warp and weft analysis, the categories still did not address the variation in complexity and thinking described by participants. To address this limitation, I returned to the analysis memos to contextualize individual excerpts within the discussion of each participants' experience. The analysis memos revealed reasoning that explained why participants were applying epistemic practices. These reasons led to the integration of the categories of description into an outcome space.

4.6.4. Are We Done Yet?

Green (2005) suggests the endpoint of analysis should be determined similar to how an endpoint is reached in phenomenology: when no new insights emerge from further iterations of analysis. As such, I continued iterations of analysis as outlined by Kinnunen and Simon (2012) above until no new insights emerge.

Upon reaching the endpoint, there was no member checking of the categories of description and outcome space. Member checking is discouraged in phenomenography because the categories of description and outcome space are constructed from a synthesis of participant data, so individual participants will not have sufficient knowledge of the data to fit themselves into the model or adequately critique the outcome space (Bowden, 2005). Furthermore, presenting participants with the categories of description and outcome space will introduce new content to the participant, potentially contaminating their perception of the phenomenon (Bowden, 2005).

5. RESULTS – CATEGORIES OF DESCRIPTION

A phenomenography has two principle outcomes: categories of description that describe distinctly different ways of understanding a phenomenon, and an outcome space that describes the logical relationship between each category of description (Marton & Booth, 1997). In this chapter, I present the categories of description, which describes the different ways in which participants used epistemic practices to construct arguments. I will first contextualize this discussion by describing the four principle arguments students discussed during their interviews. As part of their research proposal, each participant constructed two research questions: one developed in collaboration with instructors, and another constructed solely by the participant. These two research questions were presented as the first two arguments in each participant's project. Additionally, as part of the Biocore curriculum, students follow a research cycle, shown in Figure 5.1. As participants

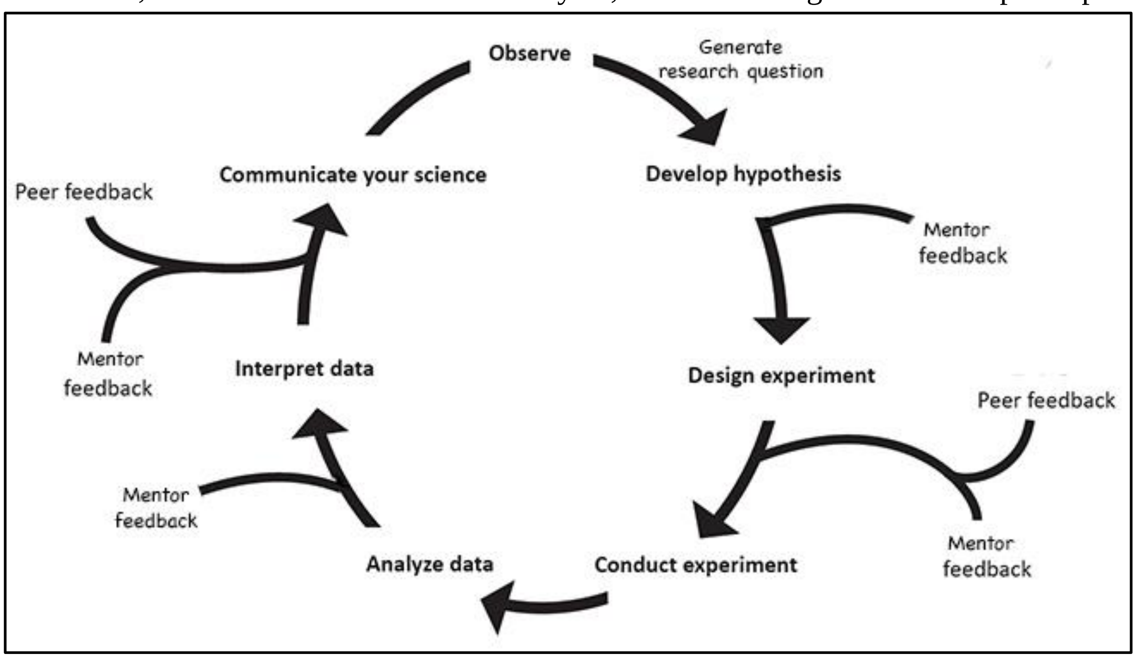

Figure 5.1. The research cycle

proceeded through the research cycle, they argued if they should reject or fail to reject their hypotheses at the analyze data step, and they argued that the conclusions they drew were correct at the interpret data step. Below, I summarize the four arguments every participant included in their research proposals.

1) The instructor-guided hypothesis was reasonable
2) The student generated hypothesis was reasonable
3) The proposed hypotheses should be either accepted or rejected
4) The conclusions drawn are correct

In describing these arguments, students discussed four different types of practices: *gathering information*, *checking credibility of information sources*, *constructing claims*, and *supporting claims*. In the following sections, I first provide an overview of the categories of description, then I present examples of how participants integrated their local vs. global perspectives and epistemic cognitive vs. epistemic metacognitive thinking with each of the epistemic practices they used to construct arguments.

5.1. Overview of the Categories of Description

Through analysis of the interview data, an interesting pattern emerged from the data. Participant discussions of epistemic practices differed along two continuums: local perspective vs. global perspective, and epistemic cognition vs. epistemic metacognition. Discussions with local perspective remained bounded within the classroom context, for example gathering information only from course content such as the laboratory manual, content the student learned in class, or literature provided by the instructor. On the other hand, gathering information from scholarly articles would be considered a global perspective since the information comes from a place outside the classroom context, and was not provided by a classroom expert such as an instructor.

Students also discussed how epistemic cognition and epistemic metacognition figured into their arguments. For example, some students used experimental evidence to argue the validity of *a statement*, which is epistemic cognition because the student is focusing on the statement itself. Other participants discussed *how certain* they were about

	Epistemic Cognition/Narrow	Epistemic Cognition/Broad	Epistemic Metacognition/Narrow	Epistemic Metacognition/Broad
Gathering Information	Gathered information from course materials, classroom experts and/or participants' own experience	Gathered information from sources outside course materials, such as peer-reviewed journal articles, or databases	Gathered information from course materials, classroom experts and/or self, but also questioned certainty of the information gathered	Gathered infromation from outside course materials, and cited own lack of knowledge on subject as a reason for seeking these sources
Checking Credibility	Trusted information because it came from course material or classroom expert	Trusted information because it comes from a database, is peer reviewed, has the correct formatting, or comes from a journal	Trusted materials provided by classroom expert because of participant's perception of what the classroom expert knows	Trusted the information becase of the way the information was gathered, or participant's perception of the author's knowledge
Constructing Claims	Used information from course materials, classroom experts, and/or self to construct a mechanistic explanation/hypothesis	Used information from sources outside classroom materials to construct a mechanistic explanation/hypothesis	Used information from class materials, classroom experts, and/or self to determine what is known or not known. Developed hypothesis/explanation to explain unknown	Used information from outside course materials to determine what is known or not known. Developed hypothesis/explanation to explain unknown
Supporting Experimental Claims	Used data and analysis from classroom experiments to support a claim	n/a	Considered how data are collected in science, and as a result, determined if data were reliable/unreliable. Supported/did not support hypothesis based on this reflection	n/a

Table 3. Summary of categories of description. The epistemic practices are presented in rows, and the ways in which participants discussed each practice are presented in columns. A brief description of each category of description is presented where each epistemic practice and the way in which the practice was described meets. Participants did not describe *supporting experimental claims* in a broad scope, therefore no category of description is presented for those sections.

the validity of the information, which is epistemic metacognition because they are thinking about *their own or others' thinking*. The definitions of these two constructs are grounded in the Epistemic Thinking framework established by Barzilai and Zohar (2014, 2016). Table 3 summarizes the categories of description resulting from the integration of student epistemic practices with epistemic cognition/metacognition and local/global perspectives.

5.2. Gathering Information

Participants discussed gathering information in all of their arguments, but the way they gathered information varied according to the arguments they were making. When arguing that the instructor-guided hypothesis was reasonable, students gathered information with a local perspective, drawing primarily from classroom resources. However, when students argued for their student-generated hypothesis, they gathered information with a global perspective, drawing information from classroom resources as well as resources from outside the classroom such as peer-reviewed journal articles, scientific databases, and books. Participants' reasons for gathering information from scholarly literature also differed. Some participants searched scholarly literature because classroom experts told them to do so, while others did so because they felt their own expertise was inadequate.

Participants who discussed the practice of gathering information described how they recalled, found, and assessed the value and relevance of the information to their project. Some students searched for information from within the classroom context. For example, Mary describes how she uses information she learned from high school and the laboratory manual to explain the instructor-guided hypothesis, a local perspective.

> *A lot of it we know… a lot of it is from our lab manual, and that kind of thing. We know that the purple color is due to 19 different genes, and that was given to us. We knew through some other research that when you cross dominant genes, you'll get more dominant genes. It's like…I don't know, we've been learning that since high school, the crosses and that kind of thing.*

Participants also gathered information from outside the classroom context, looking in scholarly journals to find prior research on their subject matter. Mary describes how she gathered information from prior literature on related plant species to explain the reasoning behind her student generated hypothesis.

> *We had to look into prior research, and to see with different plants. We did a lot of things on grapes and strawberries because they also produce anthocyanin, which is the pigment that we are studying… the purple pigment.*

The reasons for gathering information from scholarly journals varied among participants. Some, like Connie, searched for information in databases because classroom experts such as instructors or librarians told them to do so.

> *So like the TAs and during lab, especially to find literature sources... we had to sign up for a little meeting thing at the library where the librarians went over how to find good research and literature sources. I guess if they say that it's acceptable, then that's what I'm gonna do for my proposal.*

Other participants, like Karen, offered more reflective reasons for gathering information from prior research.

> *I don't know anything about these plants. They're completely new to me, so I basically have to rely on actual researchers who do know about these plants.*

The two excerpts by Connie and Karen above differ in their focus; Connie's statement focuses on the acceptability of using databases to find "good research and literature sources," while Karen's statement focuses on her lack of knowledge compared to people she calls "actual researchers." Connie's focus on obtaining good literature sources relates to obtaining accurate information, an epistemic goal. Additionally, her consideration of the information itself identifies this statement as epistemic cognition. On the other hand, Karen's statement about her (lack of) knowledge compared to "actual researchers" describes her thinking about her own cognition (what she does or does not know), which is characteristic of epistemic metacognition.

Overall, participants' search for information differed along a continuum of local scope (seeking information from within the classroom context) to global scope (outside the classroom in prior research studies published in scholarly journals). Additionally, they thought about gathering information differently, with some focusing primarily on the appropriateness and relevance of the information (epistemic cognition), while others considered how their knowledge compared to that of experts (epistemic metacognition).

5.3. Checking Credibility of Sources

While participants discussed checking the credibility of their sources through all the arguments, they discussed this practice the most while arguing for the reasonableness of the student-generated hypothesis. Participant practices ranged from trusting information because it was gathered from authorities to considering the expertise of the sources when checking credibility of information from instructors. Furthermore, participants discussed how the source constructed the knowledge, and the method of construction is connected to the credibility of the source.

Students assessed the credibility of information by considering the expertise of the source, agreement between multiple sources, and review by other experts. Some students, like Megan, considered a source credible because it was given to them by an authority, in this case, the course instructors.

> *The ones that [instructor 1] and [instructor 2] gave us, we didn't really question because our professors are giving it to us, we can assume they're pretty reasonable sources.*

Megan's perspective on checking credibility of her sources focuses locally on the classroom context. Furthermore, she considers the validity of the information itself, identifying this statement as epistemic cognition. Other students, like Bill, also trust the classroom authorities, but consider their expertise as a reason for trusting them.

> *I mean, this has been their whole career, and their whole life's work pretty much. So, I think they have more than qualified, in my books at least, to that I trust the information they're going to be giving me. That it's not going to be wrong or something.*

Like Megan, Bill's perspective remains locally focused on the classroom context (the instructors), however, in his statement, he considers the knowledge of others (the instructors' knowledge) while thinking about the credibility of this information source. As such, Bill's statement is consistent with epistemic metacognition.

When considering source expertise, participants like Hannah also took a wider perspective, considering the expertise of scholarly literature authors. Hannah trusts scholarly literature because she feels that it comes from the people who originally investigated the questions.

And so it seems like if you want to know more about the plant that you're studying, the system that you're studying it in, it makes sense to go back to the people who originally had those questions and investigated them.

Participants also considered how many times a source was cited to determine credibility. For example, Jane felt that if an article was cited by others, that there was less of a chance that the information in the article was incorrect.

I guess, it was cited by others, it was used for other people's experiments, and since so many people had been using it, I kind of assumed that it wouldn't have been corrupted, I guess, because someone would have pointed it out.

Students also considered how the knowledge was made when considering the credibility of information. Like Jane, Joe feels that others checking the information is an important indicator of credibility.

I would say that it gets back to the reproducibility part of this. If we have a lot of people that have done studies on this particular topic, and they all come to similar conclusions, the chance of that conclusion being not right, or wrong, I believe that will diminish, especially if the studies are designed so they're not all the same studies. They're differently designed but they're trying to get at the same conclusions. Then I think that the possibility that the conclusion is wrong will diminish. It may not show the big picture, it may not be complete, it may not be a complete theory, but it has stronger predictive power than if I guess something without ever doing an experiment or study.

Joe expands on the "other people checking" idea, adding that other researchers performing slightly different studies aimed at the same conclusions helps to enhance the credibility of information.

This idea of replication can also be seen in students' discussions of their own experimental data. For example, Hannah felt that her data were reliable because she had done her experiments in triplicate.

> *So like we ran in the study…we did triplicates of all the populations in all the environments so that if we could I guess be more certain that the changes that we saw were due to the selection pressures that we hypothesized, and not just due to random chance.*

Other students discussed their reasoning behind replicating their experiments. For example, Jane felt that she could not draw conclusions from her experimental data. When asked why, she explained that there were some problems with her experimental methods. She felt that performing variations of the same experiment would improve her confidence in the data.

> *I think that if we were real scientists, we would do this again. Or, like, if we were professional scientists, we would do the experiment over again, because there were a lot of thigs that, not even all our plants grew, and I don't know, we said it was because some of them didn't germinate, but I think it's 'cause we forgot to put in some of the plants. Like, I remember dropping one at some point. And then, also, we didn't put in enough plants either, so, I would wanna do several variations of the same experiment before I came to conclusions.*

There is evidence of epistemic metacognition in the above excerpt, beginning with Jane discussing how she would proceed “if we were real scientists.” Here, she is considering epistemic metacognitive knowledge about scientists, specifically how scientists construct knowledge. Additionally, she is applying an epistemic metacognitive skill (EMS) by monitoring the appropriateness of her data for drawing conclusions. At the end of the excerpt, Jane decides that she needs to replicate her experiment several times before she can draw any conclusions, consistent with her stated EMK about what reals scientists would do.

Finally, when considering the credibility of their experimental data, participants also discussed accounting for biases. For example, when explaining how she measured purple color of her plants, Mary described a procedure that her group used to ensure that the measurements were done “blind,” that is, in such a way that the raters did not know if the plants were part of the test population or the control population. She explains her reasoning below.

> *'Cause if we went in knowing that I was looking at an F2 select over red light, and I might be like, "Well, this is supposed to be darker." So I'm going to say it's the darker one even though it might not be. [...] If we went into it and just said, "Oh yeah, all the F2s's look a lot darker than the other ones." Then that doesn't actually give us a data to support it, whereas if we go in not knowing which plant we're looking at it makes it a lot harder to bias the results.*

As Mary explains, blinding the measurements helps to produce a result that is more representative of reality, a consideration of epistemic cognition. On the other hand, Hannah approaches the plant color measurements with epistemic metacognition. During the interview, Hannah states her EMK that in order for data to be considered knowledge, it must be collected objectively. She explains that since the plant color measurements are subjective, one individual may view the color differently compared to a different individual.

> *Because so the data that we collected was based on observations. It wasn't like, well we collected height data which was measured. But like the API was based on like observation. We didn't put the plants in like a spectrophotometer measure the… there was nothing that we were measuring like with a ruler or a spectrophotometer or anything like that. So because there was no measurement, it's objective [she means subjective. She corrects this later in the interview]. So a plant that I think is a four maybe to someone else is a three.*

Overall, participants considered the credibility of their information sources through both epistemic cognition and metacognition, and from local to global perspectives. Mary trusted the resources given to her in the classroom context because they came from an instructor. Bill trusts his instructors because of his EMK of the instructors' expertise in the area of his research. Jane takes a wider perspective beyond the classroom, trusting scholarly literature because it had been cited by others. Finally, Joe incorporated his EMK about how scientists make knowledge into his reason for trusting information from peer-reviewed literature.

5.4. Constructing Claims

Participants discussed claim construction at several stages of their projects. Near the beginning of their project, students talked about their transition from gathering information to hypothesis construction, their reasoning behind constructing a hypothesis, and details about how they constructed the hypothesis. Once students had performed their proposed experiments, they returned to constructing claims to state whether or not they accepted their proposed hypotheses and to draw conclusions about their results.

Students talked about using different types of information when they discussed the construction of the instructor-guided hypothesis compared to the student-generated hypothesis. For example, Hannah talks about using "foundational knowledge," knowledge she had learned in previous courses, when building the instructor-guided hypothesis.

> *Then, because we know, just like foundational science, that phenotypes that we see are usually based in both genetic and environmental factors. [...] Our reasoning was that if we could see plants that were very purple, and we knew that the other plants that they were around had been grown mostly in the same environment, we could assume that the environmental variance was just irrelevant and therefore the variations among all the plants was due only to genotypic variance. So we said, well, if we take the plants that are the most purple, then we assume those have genes that, well they have the same genes. We assume that those have alleles that code for more API [purple color]. So if we take those plants, super purple plants and breed them together, their offspring should also be more purple than if we had just taken any plant and bred it to any other plant.*

Hannah explains exactly how she moves from her classroom-focused foundational knowledge to forming her hypothesis. In the excerpt above, Hannah takes care to explain how the observable characteristics of a plant (its phenotype) can be affected by the environment and its genetic makeup. Then, using her EMK about how this information was constructed, explained how the environmental effect could be excluded from the formation of the phenotype (the plants were all grown in similar environmental conditions). Finally, she connects the parental and offspring phenotypes through alleles, which are genetic components passed from parent to offspring. Having made the connection between genetic variance and phenotype, Hannah makes the claim that breeding parents that are more purple will result in a more purple offspring population compared to breeding plants randomly.

Hanna's explanation above is an example of what the course instructors call a "Biological Rationale," or BR. The BR is a required component for each research proposal, and preceded every student's hypothesis. Megan explains what a BR is:

> *Essentially when we have projects like this, we're always told to do a biological rationale. I don't know why it's called that, never heard of it before this [curriculum], but here we go. Essentially you're supposed to start with pieces of information that you know, and have each piece of information lead into another piece of information. And at the end, it should all congeal and kind of result in your hypothesis. So it's kinda what we have here.*

Like Hannah, Nick used a BR to construct his student-generated hypothesis, however, he relied on studies outside classroom content that described nitrogen deficiency to construct his hypothesis that depriving plants of nitrogen would lead to greater anthocyanin production.

> *First, we researched effects of nitrogen deficiency on the pigment pathway itself that leads to the production of anthocyanin and we found through experiments conducted on similar species, but not Brassica rapa itself, that nitrogen deficiency ... or extreme nitrogen deficiency led to increased anthocyanin production [more purple color], which was the result of the lack of inhibitors to the anthocyanin production, which were not produced due to lack of nitrogen, which led to more anthocyanin in the offspring. By this, we concluded that both the F-2 select [population bred from more purple parents] and random F-2s [population bred from randomly selected parents], which were nitrogen deprived, would show high anthocyanin levels compared to the non-manipulated population.*

Nick builds his BR from experimental studies performed on species similar to his model system, *Brassica rapa*. Using this information, Nick describes a molecular pathway controlled by enzymatic inhibitors that connects nitrogen deprivation to an increase in anthocyanin production. Nick's construction of his student-generated hypothesis and Hanna's construction of her instructor-guided hypothesis are examples of the different ways in which students constructed claims in the contexts of different arguments.

In explaining the construction of their hypotheses, some students mentioned how they assembled information from multiple sources. Mary explains how she drew upon concepts from different sources to construct her BR:

We assembled the information on the genes and the transcript levels from one source, enzyme activity was from another one, and then stuff about the anl gene locus, that was another source, and then we had to smush it together to come up with it. I think tier two [the student-generated hypothesis] is where the majority of our literature cited came from.

The final sentence of the above excerpt suggests that information coming from research studies was primarily used to construct the student-generated hypothesis in Mary’s project. Indeed, students used course material to construct the instructor-guided hypothesis while they used scholarly journals from outside the classroom context to construct the student-generated hypothesis. This context-dependent difference in source use will be discussed in a later section.

Some students viewed using information from multiple sources as connecting between knowledge bases, an important aspect of their views about knowledge production. Dan explains this in the excerpt below.

I think the importance of connecting different bases of knowledge would be to see if you could ask questions that might be able to connect knowledge that wasn't previously known. By combining connections between all this actually did help create more knowledge and lead to more questions. The connection's a way to help generate more knowledge.

The excerpt above illustrates Dan’s EMK about how knowledge is produced. Dan feels that making connections between knowledge bases creates more knowledge and leads to more questions. These questions, in turn, lead to more knowledge generation.

Similar to Dan’s thoughts about connecting knowledge bases and generating knowledge, students like Bill see hypotheses as a means to address knowledge gaps.

> *So the, kind of a problem we ran into here was that we didn't really have any literature saying that the intervals of five degrees Celsius would have the same impact as a constant exposure to the low temperature. So, we kind of just took that as our knowledge gap. Because, I mean, there wasn't a whole lot else we could do there. I think that relates directly to the hypothesis. Like the knowledge gap should be… So the way I understand it is like; you present your whole rationale and then you say, but there's kind of, you know, something we don't know about this. This happens to be what I'm studying. And then from that you write your hypothesis and your study and all that.*

Bill determines that there was no literature supporting the idea that exposing plants to short time intervals of lower temperatures would have the same effect as constant exposure to lower temperature. This was epistemic cognition because he questions the validity of the information and finds there was no support from published literature. He then recognizes that the effect of low temperature shock was a gap in knowledge, which contributed to the construction of his hypothesis. Here, Bill applies epistemic metacognitive skills, specifically, *monitoring* certainty of his own knowledge, and upon finding uncertainty, *planning* to fill this gap of knowledge by constructing a hypothesis and testing it.

Once students constructed their hypotheses, they designed and executed a research plan to fill the knowledge gaps they had identified. Once the experiments were complete, students returned to claim construction in the form of interpreting data and drawing conclusions. While students who accepted their hypotheses saw no need to revisit their BR, students who rejected their hypothesis found themselves searching for an alternative reason for their experimental results. In the excerpt below, Cat explains her thinking once she determined that her results did not match her hypothesis.

> *What we didn't really factor in is the plants in their containers were saturated with a nutrient solution so there wasn't really partitioning of resources because the entire thing was constantly saturated. The stress induced by water loss or having to share nutrients, it wasn't there exactly. It was literally just root density and because the Brassica rapa Fast plants were bred to not mind partitioning of light and close proximity, it didn't have as much of an impact as we might have hoped.*

Cat had initially thought during hypothesis construction that placing many plants in close proximity would induce competition for nutrients via resource partitioning. However, once she performed the experiment and reflected upon the experimental methodology, she realized that the plants were not competing for nutrients. This realization contributed to Cat searching for more information, which she found in an expert.

> *One of the professors said that he tried growing a whole bunch of Brassica rapa in like a little film bottle and he could fit about 30 of them in there and they weren't doing well but they survived, so it's possible that we didn't actually cram them in enough. We thought that cramming them in four times the amount of the control should have had some impact at least but we didn't find that.*

The information from the classroom professor facilitated the construction of Cat's interpretation of her data, but she would not have found this information if she had not reflected upon her experimental methods and search for more information on the subject. Here, Cat applied an EMS, specifically *evaluating* the way the data were collected.

Other students found information to explain their unexpected results by referring to how they initially constructed their hypotheses. For example, Claire proposed that growing plants under purple lights would result in plants that were more purple in color. Once she found that her results did not support her hypothesis, she looked back at her BR and scholarly literature.

> *I think ... After we got our results, obviously it wasn't something that we were expecting, so we had to look into more literature to explain why this was happening the way it was happening. So we went back to our BR for Tier 2 [student-generated hypothesis] and looked at that pathway. And we were able to find studies that specifically looked at purple light, where in this part, when I was still creating my proposal, we weren't able to find those sources yet. By using those sources, we were able to make the comparison of the effects of purple light versus the sources we already had for blue light, and then draw conclusions why we were seeing the things we were.*

Overall, students who discussed claim construction drew from sources both within and outside the classroom context, and thought about information using both epistemic cognition and metacognition. All students explicitly stated mechanistic explanations for their hypotheses, which was likely influenced by the class requirement of reporting a BR. Interestingly, different arguments such as the difference between the instructor-guided hypothesis and the student-generated hypothesis resulted in variations in the kinds of sources students used to build their arguments.

5.5. Supporting Experimental Claims

When students discussed their experimental results, they employed practices that they did not discuss during hypothesis construction. When deciding whether or not to accept their hypotheses, students mentioned observation and statistical significance to support their decisions. For some students, observation was sufficient evidence to support their conclusions. Kbrown explains that her observations led her to reject her hypothesis because the test population had fewer high intensity purple plants than in the control population.

> *Also, just by recording the data I was looking… I was like, "Wow, there's only one in the cold environment that has an API of four, whereas in the warm environment there were multiple with an API of four." That was kind of telling.*

Students measured statistical significance by calculating the standard error and displaying the data in graphical form. They displayed the standard error in the form of error bars, and determined if differences in means were different by examining the overlap in error bars between samples. Flo explains the procedure:

> *And the way that they asked us to assess significance and our results in this class was we would graph the averages on a bar graph and then have standard error bars that were … that they had a specific multiplier based on our sample size. So I think they were around 1.5 or something, depending on the sample size and if those*

> *error bars overlapped with the next error bar, then we would know that the results weren't significant.*

The phrase "the way that they asked us to assess significance" suggests that Flo views this procedure in the classroom context. Indeed, other students discussed statistical significance as if it was an assessment of correctness. For example, Janice felt that a hypothesis could not be supported without statistical significance.

> *Well, if it's not statistically significant then you cannot support your hypothesis. And when we were doing this experiment, we really wanted to support our hypothesis so it being statistically significant is very important, otherwise we are supporting it and we have to reject our hypothesis and go over it again.*

Janice's excerpt above is an example of working from a local perspective, that is, determining specifically if her hypothesis is correct or not. There were no examples of supporting experimental claims from a global perspective, which I see as participants discussing how authors supported their experimental claims in published literature. Additionally, it focuses on the correctness of the hypothesis, which represents epistemic cognition. Other students discussed statistical significance in less concrete ways. For example, Dan agrees that non-overlapping error bars represents significance, but he has a more nuanced view of how to interpret the significance when the error bars overlap.

> *The other one we did was just the bars, the significant bars, like the percentage bars, the error bars on the graphs that we created later to see if they were overlapping or not overlapping. If they weren't overlapping, then it would show it as a significant difference, because then that would vary between there, but still be*

greater. But then if they were overlapping, that would give us a gray area where we wouldn't know for sure based on those, because they could fall in between both error bars.

Dan views significance with less certainty, presenting the overlap in error bars as ambiguous, rather than cause for an outright rejection of the hypothesis. Dan's openness to ambiguity allows him to delay closure, which may have contributed to his decision to consider his observations in addition to statistical significance when deciding to accept or reject his hypotheses. Dan explains that his decision to accept the hypothesis was based on agreement between his observations and statistical significance.

We trusted our results because when we observed the plants there just a very stark difference between populations, as well as when we did statistical tests that have shown significance, they were significant, sort of, mathematically, but also through our eyes. We were able to see that there was definitely something going on in our populations, and so that's why we trusted them, 'cause all of the population was affected. It wasn't just a couple plants, but it was everything was affected, so we were like, "There's definitely something going on here."

However, other students discussed how their perceived observations did not agree with their calculated standard error. For example, during the interview, Nick says that he believed there was a difference between the means of his test population and his control population, but then he states there was no statistical significance between the two populations. I asked him about these two statements, and in response he stated:

First of all, we concluded that due to the fact that the means were consistently higher for the groups we expected it to be higher for, so there was no inconsistencies in that statement. It was just the fact that it was not significant enough to render our hypothesis as being accepted. It was also extremely close to being accepted, but there was still some overlapping regions as I explained before, which ended up causing a systematic rejection of the hypotheses.

Nick begins his explanation by making it clear that the observed mean of the test population was higher for the test population, which does not necessarily mean that the differences were statistically significant. It is interesting that he discusses the consistency between the two statements he made. This monitoring of consistency between his thinking in the two statements is an example of an EMS, through which Nick is able to explain why both of his statements are true.

In summary, students who discussed supporting experimental claims focused on the results of their laboratory experiments. Since they were arguing for acceptance or rejection of the hypotheses generated in this course, it is unsurprising that their perspectives were locally focused in the course context. While students were focused on the validity of accepting or rejecting their hypotheses (epistemic cognition), some, like Nick, used EMS to aid in constructing their arguments. There is also evidence that need for closure described in the lay-epistemic theory (Kruglanski, 1990; Webster & Kruglanski, 1994) can affect the practices students use to support experimental claims. While Janice sought closure by viewing statistical significance as an assessment of absolute correctness, Dan viewed significance with less certainty, opting to augment his interpretations with experimental observations.

6. RESULTS – THE ANALYTIC FRAMEWORK

While the previous chapter described the categories of description (summarized in table 5.1), this chapter explains an analytic framework that I developed to understand the relationships between each category. The categories were examined first, in terms of their prevalence within participants' discussions, and second, in terms of how the prevalence changes in the context of the four arguments described in Chapter 5. These relationships will be further explored in Chapter 7 through the construction of an outcome space, which is an outcome of phenomenography that explores the logical connections between each category (Marton & Booth, 1997).

6.1. General Practice Distribution

I am visualizing the framework as four quadrants defined by two orthogonal axes, one bookended by epistemic cognition and epistemic metacognition, and the other bookended by global and local perspectives. Organizing epistemic cognition, epistemic metacognition, global and local scopes on orthogonal axes allows us to visualize the framework as four quadrants, each representing combinations of epistemic cognition or metacognition and global or local perspectives. Placing each of the four practices (gathering information, checking credibility, constructing claims, and supporting experimental claims) into the quadrants helps not only represent their occurrence, but also *how* students discussed each of these practices (Figure 6.1). Each practice is presented as a bubble whose size represents its prevalence in each quadrant as opposed to simply reporting its presence or absence. The relative positions of the bubbles in each quadrants have no meaning. This representation presents an opportunity to describe how students discussed these practices in general, and, in later sections, how their discussions of these practices changed in response to changes in context.

I determined the prevalence of the four practices in each quadrant by reading each transcript to establish whether or not each category of description (Table 3, Chapter 5) was discussed by each participant. If a category was represented, I marked it with a 1, no matter how many times the participant discussed this category within the transcript. If the category

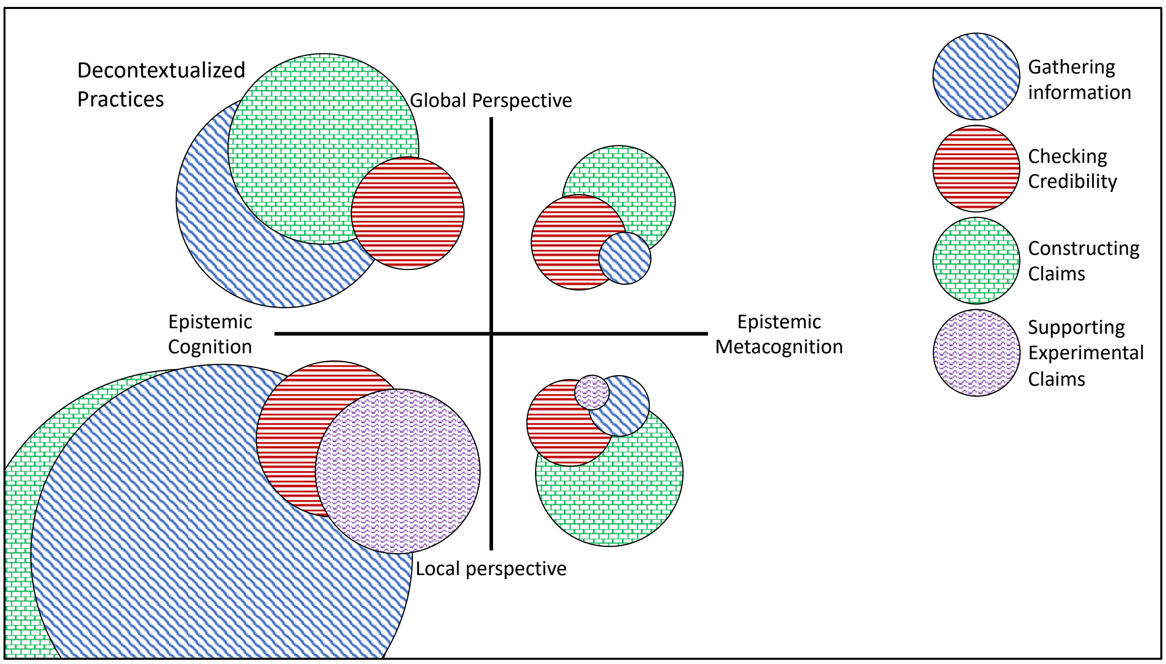

Figure 6.1. General Analytic Framework. The outcome space is organized along two orthogonal axes: one representing a continuum between local and global scope, and the other representing a continuum between epistemic cognition and metacognition, resulting in four quadrants. The four practices of *gathering information, checking credibility, constructing claims,* and *supporting experimental claims* are presented as bubbles whose size represents the prevalence within each quadrant. The position and superposition of the bubbles within each quadrant hold no significance. Participant discussions about *gathering information* and *constructing claims* were most prevalent overall, and participant discussions most often fell into the epistemic cognition/local perspective quadrants. Participants only discussed *supporting experimental claims* in the local perspective quadrants.

was not represented, I marked it as 0. I measured prevalence in this way to account for differences in the amount of participant commentary, and to ensure that the number of times participants discussed a practice was not misconstrued as a measure of the importance of the practice.

It is important to note that the data represent what participants *discussed*, not what they *thought*. As such, the absence of a practice within a quadrant means only that the participants did not discuss the practice in such a way. I cannot comment on whether or not the student *thought* about these practices in particular ways.

A general overview reveals that students discussed the practices of *gathering information* and *constructing claims* the most. Additionally, their discussions of all practices were most prevalent in the epistemic cognition/local perspective quadrant. In this quadrant, students discussed validating their ideas using a narrow range of resources such as instructors, lab manuals, and other instructor-provided, classroom-based resources. Students discussed less often how the ideas were formed or how they or others ensured that the ideas were valid, which is reflected in the outcome space.

Participants discussed the practices of gathering information, checking credibility, and constructing claims across all four quadrants, but only discussed supporting experimental claims locally. It is interesting that while participants drew from peer-reviewed literature that reported experimental results, they did not discuss how that literature supported the authors' experimental claims. Instead, students focused on the experiments they performed during their research projects, so their comments remained within the boundaries of the classroom context.

Students discussed their practices of gathering information and constructing claims the most, which is unsurprising since their project was a research proposal and report. Both practices were more prevalent on the epistemic cognition side of the outcome space. During interviews, when participants were asked how they came up with their ideas, they reported information gathered from sources, and presented it as a mechanistic explanation. Some participants elaborated more on their own certainty about the information to explain how their hypotheses came to be. These elaborations account for information gathering and constructing claims practices in the epistemic metacognition side of the outcome space.

Discussions about checking credibility were less prevalent than constructing claims in every quadrant, but were equally distributed across the outcome space. However, checking credibility was not equally distributed between the context of the four arguments described in Chapter 5. Checking credibility was more prevalent when students were discussing their student-generated hypotheses than in any other argument. This context-dependent practice usage provides important insights into the framework, as discussed below. In the next four sections, I will describe how students discussed their epistemic practices in the context of each of the arguments students presented:

1) The instructor-guided hypothesis was reasonable
2) The student-generated hypothesis was reasonable
3) The proposed hypotheses should be either accepted or rejected
4) The conclusions drawn are correct

Each argument's relative position on the research cycle is indicated on their corresponding analytic frame.

6.2. Argument 1: Instructor-guided hypothesis

The instructor-directed hypothesis was generated through classroom discussions of basic hereditary principles, past classroom data, and instructor provided resources such as publications and expert testimony. eighteen out of twenty-one participants presented the same hypothesis as a result of these discussions: that breeding parental *B. rapa* plants with 20% highest anthocyanin production (API) would produce an offspring population that was skewed toward higher API as compared to a randomly bred control population. The other three students presented the hypothesis that API operated according to Mendelian genetics.

A cursory examination of the outcome space for the instructor-guided argument (Figure 6.2) shows a vastly different distribution of practices as compared to the general outcome space. Student discussions are skewed heavily to the epistemic cognition/local

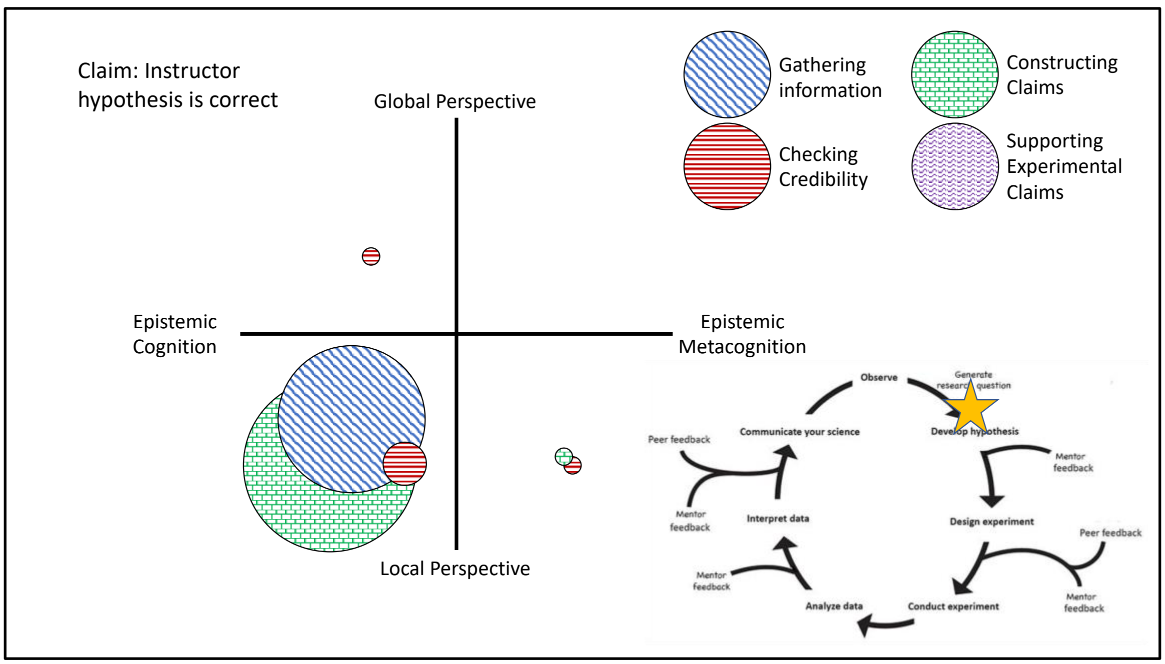

Figure 6.2. Participants' use of epistemic practices while arguing the instructor-guided hypothesis is reasonable. Compared to the general analytic framework, participants overwhelmingly discussed *gathering information* and *constructing claims* in the epistemic cognition/local quadrant. There were few discussions about credibility, and students spoke of information they gathered as fact.

quadrant, with pockets of discussion in epistemic cognition/global and epistemic metacognition/local. No students discussed the instructor-guided argument globally and with epistemic metacognition.

Student discussions for this argument centered around gathering information from classroom materials to support information participants had already learned in this and other coursework. Students labeled this information "foundational science" and did not discuss the credibility of the information. Hannah explains what she means by foundational science.

> *Yeah so I guess just like the, I don't know I guess it kind of goes back to like the central dogma. So like we have DNA, that DNA codes for RNA, that RNA is then translated into proteins and those proteins express themselves in various ways in the body depending on what they do. So we know that the purple is due to a protein, therefore the protein must come from RNA which must come from DNA, which we call a gene.*

The central dogma of molecular biology describes how information flows within a cell, and is taught as part of introductory biology curriculum. Hannah uses the central dogma above as fact to make the connection between heredity of genes and heredity of the purple color phenotype. Participants used foundational science much like Hannah did in her reasoning, treating the information as known fact.

Students also discussed gathering information from the laboratory manual and instructors. While some students did not discuss the credibility of these two sources, others said they trusted the lab manual because it was written by the instructors. As such, these statements were categorized as local because it pertained to course materials, and as epistemic cognition because its focus was on the validity of the information. However, two students felt that the laboratory manual was credible because it cited research articles. Since the research articles originated from outside the classroom environment, these discussions were placed in the global quadrant.

Students also gathered information from instructors, trusting the information because it came from people who were perceived experts and had control of participants' grades. Participants elaborated on their trust of the instructors, commenting on the instructors' scientific experience and knowledge about scientific concepts. Since these comments indicate that trusting instructors was based upon students' perceptions of the instructors' knowledge, the comments were placed in the local/epistemic metacognition quadrant.

6.3. Argument 2: Student-generated hypothesis

The second hypothesis in each students' project stemmed from an original research question generated from a student's topic of interest. Students searched for information using databases to find scientific literature on their topics, resulting in a diversity of hypotheses. Students were limited to measuring the API of plants, so student hypotheses revolved around the anthocyanin production pathway. Nonetheless, students felt that the two hypotheses were distinct from one another. While the instructor-directed (tier 1) hypothesis investigated something that was known, the student-generated (tier 2) hypothesis investigated something unknown, as Mary explains:

Then tier two was we came up with what we wanted to do, so we said red and white light and we had to come up with a rationale for that. The tier one rationale was hidden within the lab manual, but we just had to visually show it whereas tier two we actually had to dive into the research and figure it out for ourselves.

This change in context comes alongside a change in the distribution of practices in the outcome space (Figure 6.3). The dominant practices shift from the epistemic cognition/local quadrant to the epistemic cognition/global quadrant. This change is facilitated by students' shift in discussion from gathering information from course materials to gathering information from scientific literature. Despite a change in scope, gathering information remains anchored in the epistemic cognition quadrants.

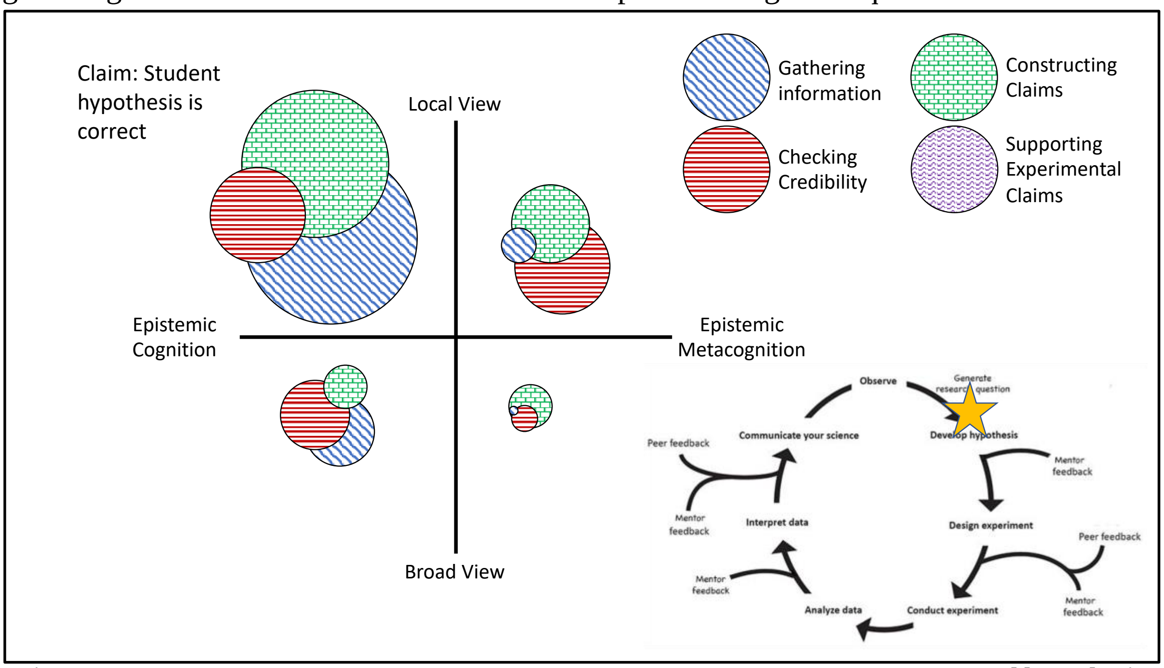

Figure 6.3. Participants' use of epistemic practices while arguing their student-generated hypothesis was reasonable. Compared to the instructor-guided hypothesis, participant discussed *checking credibility* more often. Additionally, the frequency of participant discussions decreased in the epistemic cognition/local perspective quadrant and increased in all other quadrants.

Checking credibility becomes comparatively more prominent in the student-generated argument outcome space, occupying all four quadrants. In this less certain context, students began to discuss how they determined if information was reliable. In the local perspective, participants still trusted the expertise of instructors and course materials. Likewise, students trusted materials from scientific literature because of the perceived expertise of the study authors. Additionally, knowing that articles were found on a database, were peer reviewed, or had the correct formatting were sufficient for some students to consider a source reliable. Since students used each of these criteria to determine the credibility of a source, these practices are rooted in epistemic cognition. However, some students also provided further reasoning for considering a source reliable, commenting on how the information in the scholarly articles was produced. For example, Ella comments on how she perceives peer review, and why this method of producing information helps with the accuracy of the information.

> *Because it's not just one person saying this. They could be saying an outrageous thing and nobody asks ... I don't know what I'm trying to say. […] I think because there's more opinions going into it and more views. To me, it seems like whoever's peer reviewing it is trying to find everything that is wrong with it. It's just taking an idea and critiquing it. It should be pretty accurate, I would assume.*

Here, Ella goes beyond *it is reliable because it is peer reviewed*, shifting the focus from the *outcome* of reliable information to the *process* of how the information was produced. Namely, that the article is built from more than one voice, and that critique through peer review is part of the construction of this information.

Like gathering information, constructing claims is most prevalent in the epistemic cognition/global quadrant. However, constructing claims also appears in the two epistemic metacognition quadrants. Since the assignment pressed students to address a knowledge gap for the student generated hypothesis, some students monitored what they did and did not know (an EMS) about a topic while building their hypotheses. Joe gives an example of this kind of discussion:

> *The part of a mechanism in which my group, we inferred what was happening, was that we thought that the introduction of BPA would cause more anthocyanin production than if there was no BPA. That's the part we put in our hypothesis. That's the part we don't have anything to support, but given that we have studies that show that anthocyanin helps prevent damage from oxidative stress, and we have studies that show that BPA does introduce oxidative stress on plants, then we just put those together and made our hypothesis.*

In the beginning of the excerpt above, Joe explicitly states that BPA causing an increase in API was an inference, a claim that did not have data directly supporting it. He explains that this uncertainty is the reason this claim became his hypothesis.

The shift in context between the instructor-guided and student-generated arguments was accompanied by a shift in the prevalence of practices in each quadrant of the outcome space. The student-generated argument was a less certain context that compelled students to seek information outside course materials, and have more discussions about credibility. In this context, students kept their discussions primarily in the epistemic cognition space, but also talked about practices that were in the epistemic metacognition space more in comparison to the instructor-guided argument. In thinking about credibility, some students began to talk about the knowledge production process as a part of their credibility assessments. These epistemic metacognitive discussions facilitated reflection about how knowledge is produced in biology. Likewise, some students monitored what they did and did not know in the context of the second argument, where they were asked to construct a hypothesis that addressed a knowledge gap.

6.4. Argument 3: Reject or fail to reject hypotheses

Once students constructed their hypotheses, they proposed an experimental plan to test their hypotheses. I had expected the validity of the experimental plan to be an argument, but students universally listed their plans as a statement rather than an argument (*i.e.* "This is what I am going to do…" as opposed to "This is an appropriate way to

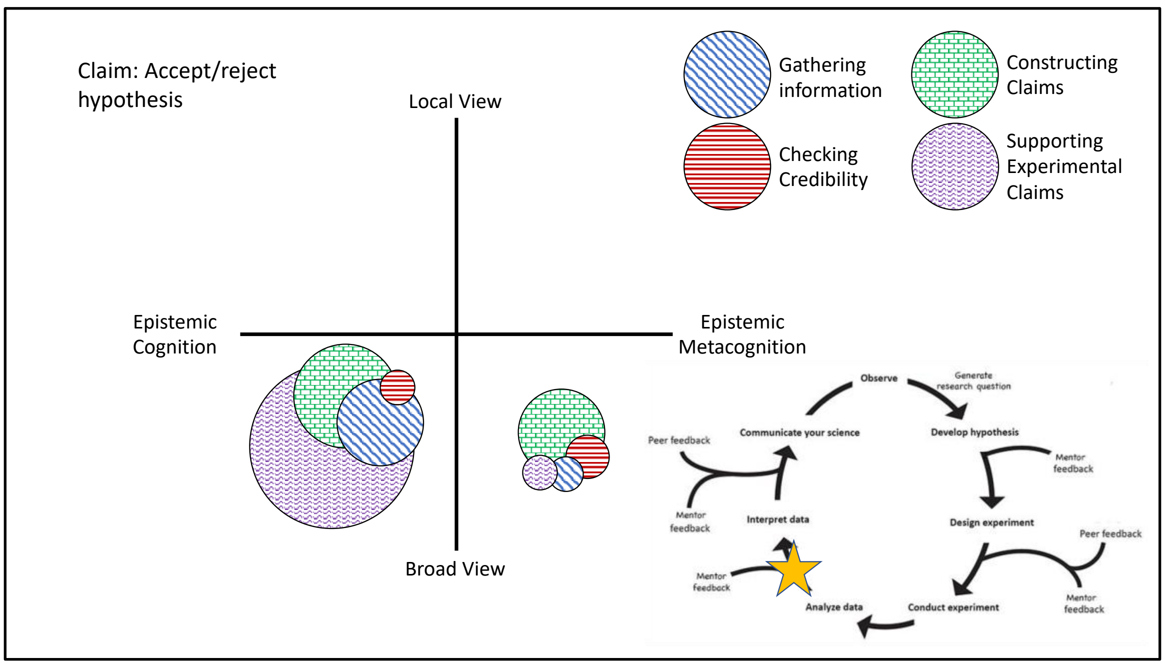

Figure 6.4. Participants' use of epistemic practices when arguing their proposed hypotheses should be accepted or rejected. While discussing this argument, participants only discussed their epistemic practices in the local quadrants. Additionally, this argument was the only context in which participants discussed *supporting experimental claims*.

investigate because…"). As such, the next argument participants made in their projects was if they decided to reject or fail to reject their hypotheses.

Gathering data for this argument consisted of collecting and analyzing data from experiments that participants performed. Students measured anthocyanin production by observing the purple pigmentation that resulted from the accumulation of anthocyanin in plant tissue. Students obscured the identity of test and control plant samples before measuring color against a color standard provided by the instructors. Participants felt that this was an appropriate method to measure plant color because the instructors had told them to do so. Some participants elaborated on their reasoning, saying that deidentifying the plants helped to eliminate bias, aligning their data collection methods with their EMK that science knowledge is objective. Since students concentrated on experiments they performed in the class, the gathering information practice appears in both epistemic cognition and epistemic metacognition quadrants of the local perspective.

Students used the data they gathered from their experiments to support their experimental claims of accepting or rejecting hypotheses. Participants discussed two distinct forms of support: observation and statistical significance. As discussed in chapter 5, participants discussed observation and statistical significance both separately and in combination to support their claims, discussing them both cognitively and metacognitively. It is interesting to note that while participants discussed observation and statistical significance as support for their own experimental claims, they did not discuss these items while discussing the instructor-guided or student-generated arguments, where they presumably used data that were likewise generated from experimental methods.

While there were few discussions about credibility in the context of hypothesis accept/reject argument, some students did check credibility of their data by assessing how they measured anthocyanin production and how they analyzed their data. Discussions about credibility centered around expertise and objectivity. For example, Bill describes his uncertainty about his own data based on assessment of his expertise and the way they collected the anthocyanin production data.

> *I don't think myself or any of my peers have really enough experience, or really even the qualifications, to be honest, to call it scientific knowledge. I feel like there's stuff that, you know, not even methods wised, but like others just the rationale and the thinking. There could have been something that we missed. That I guess that could have confounded the results. Also it's kind of, I mean this is a largely qualitative experiment. We were assigning numbers to, I don't know, I guess the numbers make it quantitative but still. Like we were literally just looking at plant colors with a card and saying, yeah it looks like a one. You know? Whereas if we would have thrown up like a spectrometer or something and used actual like, data measurements like that, maybe it would have been different. I'm not sure.*

In the above excerpt, Bill uses his EMS (Epistemic Metacognitive Skill) to monitor and reflect upon the uncertainty about the data he collected. He struggles with his own lack of expertise, questioning whether or not he had though through all the possibilities. Additionally, he reflects on qualitative and quantitative data, and how they fit within scientific knowledge. His discussion of these topics was made possible by the context of the hypothesis accept/reject argument, where students had performed an experiment with their own hands.

6.5. Argument 4: Interpretations and Conclusions

Argument 4 is closely related to the accepting or rejecting hypotheses argument presented above. Once students decided to accept or reject the hypothesis, they made claims about the reasoning behind whether the hypothesis was accepted or rejected. Students that failed to reject their hypothesis did not discuss argument 4 as deeply because they referred to the BR, which they used to rationalize their hypotheses. Students who rejected their hypotheses returned to the information gathering practices employed for instructor-guided and student-directed arguments such as using course materials and peer-

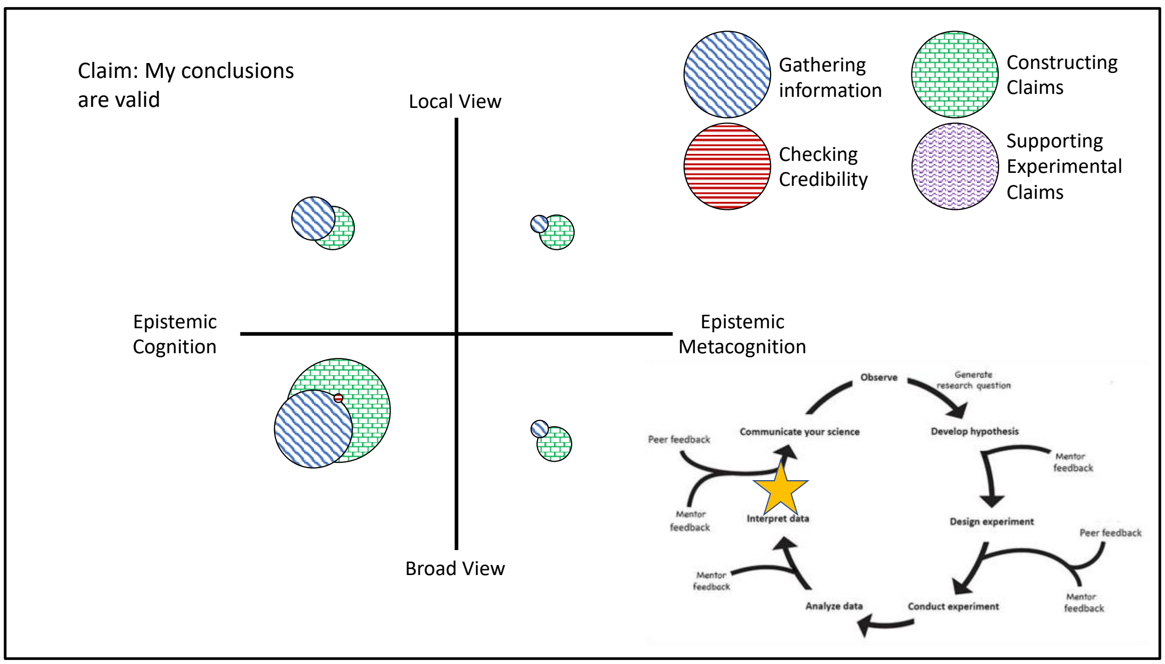

Figure 6.5. Participants' use of epistemic practices when arguing the conclusions they have drawn are correct. Participants did not discuss this argument as much as the other three arguments, as evidenced by the small size of the bubbles. While discussing this argument, participants talked almost exclusively about gathering information and constructing claims.

reviewed literature as information sources. The outcome space for the conclusions argument is dominated by gathering information and constructing claims practices, with checking credibility almost completely absent. One reason for the few discussions on credibility could be that discussion of students' experimental results occurred at the end of the interview, when participants had already discussed checking credibility of course materials and peer-reviewed literature.

The ways that students constructed claims in the conclusions argument were influenced by students' previous arguments, particularly information gathered to support those arguments. Participants discussed if the conclusions they had drawn from the instructor-guided and student-generated arguments were correct, and referred back to information used to construct the arguments as a starting point to construct claims about the experimental results they had obtained. For example, after finding unexpected results from her experiment, Ella searched for reasons within her BR.

> *I think it's because the biological rationale is still there, that's not going away, so based off of that ... I feel like this is circular reasoning, but, because we were expecting this based off of our logic, we're still assuming that that's true, but we just didn't see it. Then we were thinking how could we see it if it's still true, and then that made us think more generations ... Because then slow changes can lead to ... It's slowly changing every time, but just not significantly. Then if you go back to the first one, the first generation and the last generation and look at those two, then you should see a difference.*

Ella seems to be weighing the strength of her BR compared to the strength of her experimental data. In the end her logic wins out, and she searches for a reason why her BR could still be accurate while maintaining her results. In this case, Ella uses uncertainty about her results ("we just didn't see it") to reconcile the disagreement between her results and her BR.

Constructing claims in the conclusions argument is also linked to the contexts of the instructor-guided and student-generated arguments. For example, Bill rejects his hypotheses for both the instructor-guided and student-generated arguments, but interprets the results differently. For the instructor-guided argument, he sees that the data do not support the hypothesis but argues the hypothesis may still be true.

> *But yeah, basically, I do think that, I think that even though you couldn't see it from, because the discreet gene was maybe masking it, if you want to call it that. I think that the allelic variation in the select population was lower and was more shifted towards the more alleles that would produce more anthocyanin pigment. Again, I just base that on what I've learned in class and I guess, how genetics seems to work.*

He bases this conclusion on what he learned in class, his understanding of how genetics works. Consistent with the outcome space for the instructor-guided argument, Bill bases his reasoning on information from within the classroom context. However, when he talks about the student-generated argument results, he is willing to let go of his hypothesis. Below he gives his reasoning for why he does not believe the cold shifted the plant API.

> *That was mainly from our data. I mean, the means, the actual number we calculated for the means were different. But the standard error bars overlapped. So I guess it*

could have been possible that they were shifted, but we don't know from our sample size and from our data. I think, I don't know if lower temperatures would have been too low for the plants, or would have killed them. Or if they would have been, or if 12 hour intervals would have been better. That was kind of the gray area. Why we chose the eight hour intervals. Again, I think we were just trying to mimic the idea of a night cycle. But we really didn't have any study that backed eight hour intervals specifically so we were kind of just spit-balling there. So I don't know if a 12 hour or a 16 hour would have been better. You know, or if we should have continued it for two weeks instead of just five days.

What is interesting about the excerpt above is Bill's reflection on the certainty of his results. He lists many uncertainties such as "Were the temps sufficient?" "Were the interval times ok?" and "We didn't have a study backing our methodology." These kinds of questions were not raised in his reasoning for his first argument. When I asked Bill why he thought differently between his two arguments, he explains:

Well, when we went into this, we, I think we talked to [professor], the gentleman that breeds the plants. And we were talking with [instructor 2] and [instructor 1] and they all said, no one in the lab has done this experiment before. So it will be interesting to see how it turns out. So, I guess we didn't really know what to expect. And we haven't done anything about biosynthesis pathways or cold effects on plants in class. And I don't have any background knowledge about that really, aside from what I learned in the papers.

So, yeah I guess that's part of it. But, I think also that added environmental variance. Which knowing from my, the first paper we did, environmental variance is like, there can be a ton of stuff going on there. So I wouldn't really be surprised if there is something else that was keeping them from expressing more pigment. So, I suppose that there is a lot of genetic variance as well. But, I guess the way ... I don't know. They seem so certain of it in lecture, when they explain it, like, this is the way it is. Like it's, you know. I don't know. Maybe I'm putting too much faith in that.

His explanation points to a few things. First, the uncertainty of the second argument plays a role: no one had done this before, no one knew what to expect, and there were holes in his background knowledge. On the other hand, with the instructor-guided argument, he has experience with confounding factors from his first project, especially from environmental factors. Additionally, his perceived certainty in which the material was presented in class also had an effect on his decision. It seems that he views instructor-guided and student-generated arguments through two different epistemic frames: one certain and the other uncertain.

6.6. Connections Between the Categories

Tracking how participants used different epistemic practices between arguments reveals some of the logical connections between the quadrants of the analytic framework. Comparing the outcome spaces of the instructor-guided and student-generated arguments provides us with some ideas about the connection between local and global perspectives because of the shift in practice prevalence from local to global perspective. Discussions about the instructor-guided argument centered on classroom goals, such as aligning knowledge with the instructor. However, discussions about the student-generated argument aligned with scientific goals, for example gathering information to explain a phenomenon. These differing goals are reflected in Bill's interpretation of his experimental results presented above. Since argument 1 was instructor directed, it had a "correct answer," which Bill had deduced from his coursework. When his results disagreed with his perceived correct answer, Bill defaulted back to his classroom goal to align his knowledge with the instructors. However, Bill's interpretation of his second hypothesis experimental results aligned more with scientific goals. In explaining his interpretation, Bill asks many questions about confounding factors in his research. Answering these questions would bring him toward the goal of accurately describing a phenomenon: a scientific endeavor.

Moving between argument contexts also affected the amount of certainty with which students viewed information. For example, students felt more certain about materials they were given in class, especially since the information matched curricula that were delivered to them in certain terms during lecture. When discussions moved from a certain context, such as the instructor-guided argument to a less certain context, like the student-generated argument I observed an increase in practices related to checking credibility. These discussions about credibility were distributed throughout the framework, but an interesting trend that I observed was that credibility discussions within the epistemic cognition quadrants focused on the outcome of credibility, whereas discussions within the epistemic metacognition quadrants focused on the process of making information credible.

Additionally, students that discussed practices metacognitively generally did so after they had discussed the same practice cognitively. As such, engaging students in conversations about the process of constructing information could engagement in epistemic metacognition.

At this point I would like to clarify that this framework is not meant to be a hierarchy with global perspective and epistemic metacognition as the preferred practices. Indeed, appropriate use of practices from each quadrant specific to the context would be ideal. For example, the reason participants moved from a certain to uncertain context was because the instructors required it in the classroom. Students' goals to attain a good grade led them to pursue goals that were aligned with scientific practices such as seeking information from peer-reviewed literature. While the framework presented here illustrates the changes in epistemic practices between contexts, it only gives a surface view of participants' use of epistemic practices in context.

In the following chapter, I use the framework described above to take a deeper look at why students used specific practices in certain contexts by building composites of student epistemic practices. The description of these composites combined with a discussion of the relationships between these composites comprises the outcome space.

7. RESULTS – Outcome Space

In this chapter, I will describe three composites, representing pooled practices drawn from participants' descriptions. These composites were built around three exemplar cases: Connie, who bases many of her epistemic choices on instructor, TA and peer input; Joe, who applies epistemic practices based on his EMK about how science produces knowledge; and Cat, who applies epistemic practices for a mixture of reasons based on both classroom and scientific practice. While analyzing transcripts to build the three composites mentioned above, a fourth hypothesized composite emerged around the participant Hannah, whose conflicted beliefs based on her religion and scientific education influence how she thinks about scientific knowledge.

Each of the exemplar cases is contextualized within the research process cycle, a science process that comprises part of this CURE curriculum. A summary of the research cycle is presented in Figure 7.1. Each of the four arguments described in previous chapters can be mapped to different parts of the research cycle, and as such will act as temporal anchors for each of the cases. Since

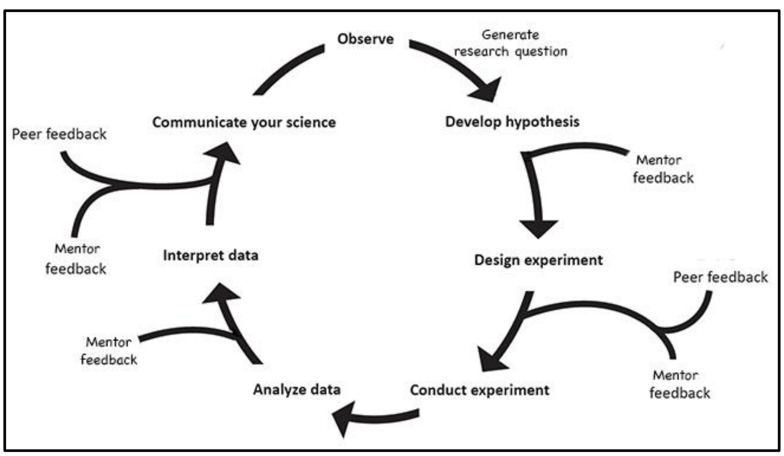

Figure 7.1. The research cycle. This figure is a representation of the research process this CURE curriculum follows. Students are expected to make observations, develop hypotheses, and design experiments to collect, analyze, and communicate data and findings. Mentor and peer feedback is provided throughout the research process.

these data were generated from participants recalling their experiences over the course of their project, these cases cannot be considered longitudinal. In these cases, I use the word "temporal" not as a passage of time, but as a marker for participants' progress in the research cycle. With respect to the four arguments, arguments for hypothesis 1 and 2 map to the "develop hypothesis" stage, arguments for accepting or rejecting hypotheses map to the "conduct experiment" and "analyze data" part of the cycle, and arguments for conclusions map to the "interpret data" part of the cycle.

Throughout this chapter, I will name and give a brief explanation of each composite, followed by a description of an exemplar for each composite in the form of a case. Each exemplar case is bounded by the individual participant's interviews and project posters to further contextualize these practices in the classroom in which they were used. The exemplar cases primarily apply practices characteristic of one composite, but they do use some practices from other composites. To conclude this chapter, I will explain the relationships between the composites as part of a complete outcome space.

7.1. Composite 1: Validating

I named the first composite Validating, since the practices within this composite are aimed at validating information by matching information and practices to what participants perceived as instructor expectations. Participants who used Validating practices also trusted information because it came from an expert, and did not discuss how the information was constructed. Therefore, practices within the Validating composite are distributed between local and global perspectives, but remain in the epistemic cognition quadrants of the Nested Argumentation Practices Analytic Framework described in Chapter 6. Given that practices in the Validating composite were used to gather trusted information from experts or match perceived instructor expectations, these practices were often social, with the participant acting as the recipient of information. The majority of Connie's practices fell into the Validating composite, so I use her as an exemplar to describe in more detail the practices within this composite.

7.2 Case 1 Connie: "I guess if they say it's acceptable, then that's what I'm going to do"

Of the three cases presented in this chapter, Connie is the participant whose practices are most rooted in the local perspective. As such, she looks to employ skills that improve her performance in the classroom. For example, when I asked Connie what importance writing this proposal had for her, Connie kept her answer in the classroom context.

> *I think for this one specifically, it teaches us how to create a proposal poster and I think that it's important for the next few years that we have in [curriculum] and for … Right now we're working on creating a grant proposal. So I think both of these proposals led us to that point, and its different varieties or the first one we had was just a paper, and this one is like a poster, and then our third one's gonna be for a grant.*

Connie believes that her experience writing this proposal is a skill that she can use in future courses in the curriculum, which underscores her desire to do well in the course. Consistent with this goal, Connie looks to instructors, TAs, and peers to help her revise her hypotheses.

> *So since we're doing a genetics proposal, because our presentation, we focused more on the surface, whereas they wanted us to go deeper in with the enzymes and the photoreceptors, and stuff like that […] so for both proposals, we do like informal feedback presentation in front of our discussion class. So it's the TAs that are there, the students in my lab and the [curriculum] directors. And so we go over our experiment and our introduction, or methods, results and everything, and they give us feedback to fix our proposal, I guess.*

The two statements "…they wanted us to go deeper…" and "…they give us feedback to fix our proposals…" show that Connie believes there is an outcome that can be judged as correct or incorrect. For the outcome to be judged as correct, she sets certain criteria to match what she perceives as what the instructors want; in this case, a deeper explanation that includes discussions about enzymes and photoreceptors. These criteria help Connie to set guidelines for what acceptable knowledge is in the classroom.

7.2.1. Instructor-guided hypothesis

Connie employs the matching criteria practice in her construction of the first hypothesis. She finds information in the lab manual, resolving to discuss phenotypes and mendelian genetics in her explanation.

Well, we spent a lot of time in our [curriculum] manual, it talks about the quantitative traits and we have been talking about genetics and the phenotypes, and all the alleles in our lecture class and during lab, and all that. So I think a rationale, it was just from a Mendel's, Mendelian trait, kind of situation. So using that, we just figured out how the next generations would look like.

Reporting information seems to be Connie's way of showing her knowledge. For example, she uses the biology disciplinary language to explain her reasoning behind her first hypothesis during the interview. However, while she reports information that is related to the hypothesis, she does not adequately connect the information to explain the reasoning behind the hypothesis.

So the [instructor-guided hypothesis] was up here and so this was more the genotypic variations and phenotypic variations. So the environmental variations is not in play. And so, we were given the F2 generation, so we had from the F1 generation, they were all heterozygous with heterozygous alleles, and the 20% most pigmented were selected. And then 20% random within the population was also selected and that's where you get your F2 generation. So for the F1 generation, because they're all heterozygous, that means that there shouldn't be any that are recessive. So that's why the curve ends right before that. So they should all show at least heterozygous alleles or dominant. And then with the 20%, most pigmented being selected, that's where you get your F2 selected. So those we're hypothesizing that the selected will be the more pigmented than the random, and we're thinking that the random would be like this little bar up here. The F2 random would be *throughout that, so it could be any of those, whereas the F2 selected, we hypothesized that it should be more pigmented.*

Connie includes all the topics required to explain why the hypothesis is viable: genotypic/environmental/phenotypic variations, parental/offspring genotypes, F1 (parents) and F2 (offspring) generations, the control population (F2 random), the test population (F2 selected), and selection of the 20% most pigmented parents. While Connie describes these topics, she does not connect them to her prediction that the F2 random population would have lower pigmentation than the F2 selected population. For example, genetic theory states that phenotypic variation (variations in the observable characteristics of an organism) is influenced by two factors: genotypic variation (variations in the genetic makeup of organisms), and environmental variation. Since all plants in the experiment are grown in the same environment, we can assume that any variation in phenotype is due to variations in genotype. This reductionism is typical of science epistemology: remove as many confounding factors as possible through experimental process so that viable conclusions can be drawn. As an individual outside Connie's mind, I am not able to judge whether or not Connie understands the role of reductionism in this experiment. What is apparent is that her focus is on the outcome of the process, that "the environmental variations is not in play" rather than the process itself, the reductionism that helps her make conclusions about genotypic variation and its effects on phenotypic variation.

Connie continues her explanation by discussing the heterozygous phenotypes of the F1 generation. She does not connect the discussion of the heterozygous genotypes to her hypothesis. Her earlier statement "So I think a rationale, it was just from a Mendel's, Mendelian trait, kind of situation" suggests that she included this discussion to demonstrate her understanding of mendelian genetics, rather than to help her explain the hypothesis.

Connie concludes her explanation by discussing the selection of the 20% most pigmented parents to generate a population of offspring (F2 selected). Connie continues to apply the practice of reporting information. It becomes clear that she does not understand how selecting the 20% most pigmented parents will result in a more pigmented offspring population when we examine the figure Connie presented in her poster (Figure 7.2). To test the hypothesis that breeding plants selected from the most pigmented 20% of the parental generation will result in more pigmented offspring compared to randomly selected parents, two offspring populations should be generated: one from the most pigmented parents, and the other from randomly selected parents. Thus, a figure depicting the populations in this experiment should display three population curves: parental, F2 selected, and F2 random. Instead, Connie provides only two population curves, one for the parental and another for the offspring generations. She represents the F2 random and F2 selected populations on the same curve, incorrectly indicating that the F2 selected population was in the hashed part of the F2 population curve. In reality, the F2 selected population curve should have been separate, and entirely shifted towards the right compared to the parental and F2 random populations. Connie's

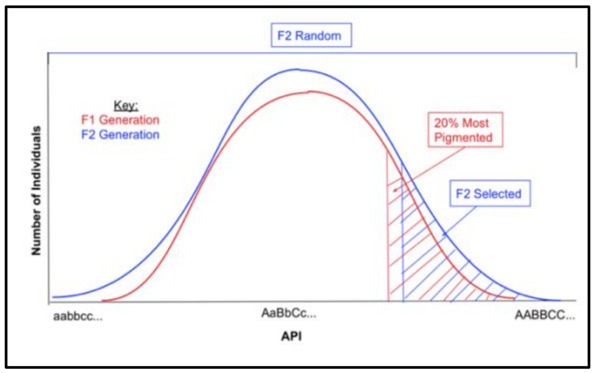

Figure 7.2. Connie's representation for populations of F1 and F2 generations. This figure incorrectly presents the F2 selected population as a part of the F2 random population. Inaccuracies in this figure paired with other practices such as listing facts suggest that Connie uses the epistemic practice of reporting information to build the instructor-guided hypothesis.

descriptions above show the limitations of matching criteria to classroom expectations. Connie includes many of the expected information in these descriptions, but without understanding the importance of connecting genetic theory to her discussions, the resulting descriptions are incomplete. While Connie's matching criteria practice helps her to determine what counts as knowledge in the classroom, pairing this practice with EMK about why scientists believe these facts (i.e. connecting her descriptions to genetic theory) would help her to construct a more complete description.

7.2.2. Student-generated hypothesis

As seen in the outcome space, Connie approaches her student-generated hypothesis with a more global perspective, looking for information from peer-reviewed literature. However, even though she gathers information from sources outside the classroom, her criteria for using literature remained firmly in the classroom context.

> *I didn't use Google Scholar before this but I was just told that usually like the UW Library database, there was sites and the research papers, and the literature sources on there are good sources to use as well as Google Scholar. We worked on it during our discussions. So like the TAs and during lab, especially to find literature sources. And then we also had to sign up for a little meeting thing at the library where the librarians went over how to find good research and literature sources. I guess if they say that it's acceptable, then that's what I'm gonna do for my proposal.*

The reason Connie looks for information in these databases was because of perceived instructor expectations. She worked with TAs to find literature, and states explicitly that if the classroom experts find it acceptable, that is what she is going to do. This practice is consistent with her goal of aligning her practices with what she perceives as the expectations of classroom experts, and in doing so, succeeding in the class.

Just as she had done in the instructor-directed hypothesis, Connie lists facts to explain her reasoning behind her student-generated hypothesis that exposing plants to red light will result in an increase in anthocyanin production.

> *So I had red light and white light. So red light only has red light whereas white light contains both red, blue and green, I think. And so the light … So the plants use the photoreceptors to respond to the light, which is right here. And then down here I have the different types of photoreceptors. So we had phytochrome as [struggling with words here] cryptochrome, and photochrome. I know the abbreviations. So we focused more on like PHY, because that's for the photoreceptors for that light. And so the photoreceptors will inhibit the COP1 and COP1 inhibits HY5. So if the photoreceptors inhibiting COP1, so then more HY5 is being produced, and that leads to more anthocyanin being produced. So the HY5 goes to the flavonoid pathway and then it creates a DFR and then the Anthocyanin.*

During the interview, Connie pointed at Figure 7.3 while she was explaining her hypothesis, listing parts of the figure as she went. Her speech slowed at this point, and as noted in the excerpt, she struggled with some of the words. Much like her explanation for the instructor-guided hypothesis, Connie fails to connect many of her listed facts with her hypothesis. For example, she mentions red light, but does not explain its mechanistic relationship with the hypothesized increase in anthocyanin production; something that biologists expect to see in order to validate a claim. At the end of the excerpt Connie speaks more confidently about the photoreceptor PHY and the other proteins in the pathway, COP1, and HY5. These were the very aspects that she was asked to address following a feedback presentation when she was developing her hypotheses: "they wanted us to go deeper in with the enzymes and the photoreceptors, and stuff like that." Just as she had done with the instructor-guided hypothesis, Connie ensured that her reasoning matched what she perceived as the expectations of the instructors. This is an example of how Connie's matching criteria bring her discussions closer to biology epistemic practices.

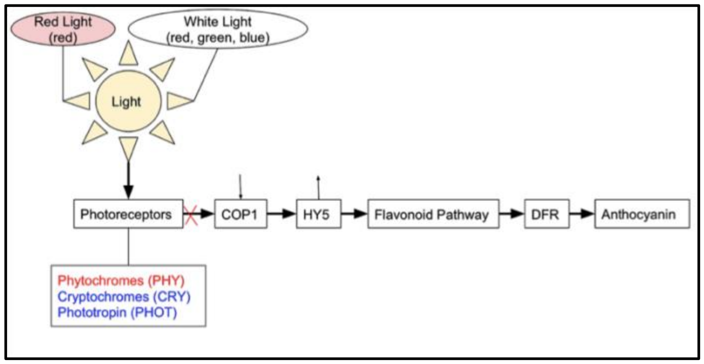

Figure 7.3. Connie's BR. Connie had difficulty explaining her BR during the interview. She seemed to be reading the text on this figure during her explanation.

7.2.3. Accepting or rejecting the hypotheses

Connie's argument supporting the rejection of both her hypotheses is the only argument in which she does not discuss seeking guidance from instructors or TAs. She describes her observations, which comprise all the support Connie mentions for her decision.

> *So, for actual results, I think we rejected both [hypotheses] of the experiment. I think the red light had less API than the white light, which is not what we were expecting. And we had pictures of our plants every two days and the red light plants had, the height was significantly more than the white, but then they eventually like averaged out the last few days, but if you compare that two, the red light, the stems were more thin and sickly looking. Whereas the other one was like very green and vibrant, and they had more leaves.*

Aside from the API for each plant, Connie also notices that the plants exposed to red light were taller at first, and by the end had fewer leaves and were more sickly. She describes how she tried to measure these two characteristics.

> *So, we wanted to talk about the height, but when we did ... we made a graph of it, there was not really precise correlation with height and API, so we couldn't focus on the height. So, we did a biomass of the plants, so cut them up and like calculated the biomass and I think the red lights had less biomass.*

Here, I see a shift in Connie's practices. Instead of trying to achieve an outcome that is judged as correct, she seems to be pursuing the measurement of an unexpected phenotype.

However, in the next section, as Connie searches for a reason behind her observations, she returns to the instructor for guidance.

7.2.4. Connie's Conclusions

A discussion with an instructor guides Connie to an explanation of her unexpected observations. The instructor suggests a reason for her observations, and Connie pursues the suggestion by searching for the definition.

> *And we like saw that the red plant was significantly higher and we talked to … I think we showed it to the, one of the program directors and [they] mentioned how it was kind of weird, but then [they] also mentioned the elongation or elotation, or whatever it's called, and so we kind of searched up the definition of what it was.*

As with her hypotheses, Connie does not seem to completely grasp the explanations she discusses. For example, she has trouble with the word "etiolation" (she calls it "elotation"), which happens when plants are either partially or completely deprived of light. This results in plants elongating to seek out light by growing taller than competing plants. Etiolation could explain the phenotype of plants grown under red light, however while Connie may understand the connection between etiolation and red light exposure, she does not discuss this in the interview.

Connie's practices change some as she proceeds through the research cycle. She shifts her information gathering practices from local (lab manual as a source) to global (published scientific literature as a source) when transitioning between the instructor guided and student-generated hypotheses. Furthermore, she shifts from relying on instructor feedback to relying on her own observations when deciding whether or not to accept her hypotheses. However, Connie reverts back to seeking instructor feedback when she is interpreting her data, and in her reasoning for seeking information from databases reveals how entrenched the instructor feedback practice is for her.

Another salient practice that was consistent throughout Connie's progress through the research cycle was information reporting. The word "because" was conspicuously absent from her reasoning behind all four arguments. As a result, her explanations presented as lists of facts rather than a connection between these facts to support her hypotheses or conclusions. In other words, Connie did not organize the data she presented into lines of reasoning.

While it was necessary to describe the limitations of Connie's reasoning as evidence for the alignment of her practices with instructor expectations, her Validating practices should not be viewed negatively. Just as Connie used her matching criteria practice to determine what counts as knowledge in the classroom, Validating practices present an opportunity for students to enter a scientific community of practice by aligning their practices with the culture of the scientific community of practice.

7.3. Composite 2: Enculturing

Much like the Validating composite, practices within the Enculturing composite were often social, but for different reasons. While Validating practices focused on matching the perceived expectations of instructors, Enculturing practices were built around information sharing within a knowledge culture. Enculturing practices included making connections between knowledge bases, connecting knowledge to society, and communicating results to others. Participants explained that they used Enculturing practices because they matched participants' perceptions of how science produces knowledge. Many of Joe's practices fit in the Enculturing composite, so I present his case as an exemplar. While Joe's practices in many ways match Connie's in the instructor-guided hypothesis argument, his practices begin to diverge as he moves through the research cycle.

7.4. Case 2 Joe: "Science isn't product minded"

Joe's motivations stretch father into the future than Connie. Whereas Connie's goals extend to future semesters of the curriculum, Joe's goals have to do with his career as a science researcher.

> *I want to go into a career doing research, so I want to go through graduate school, and then possibly a post doc, and possibly even applying for faculty positions in the future. So, making a scientific poster and actually in general improving communication skills, both being able to communicate science to fellow researchers and general public, is very important to what I will be doing in my career.*

Joe sees the project as a way to attain skills, much like Connie. However, his skill attainment is applied toward the goal of a career in science. His focus on communication skills belies his EMK about scientific knowledge.

This is not complete. Science isn't product minded. There's no end goal of science. It's a process of creating and revising knowledge about the natural world. One study is just a component of that. It serves as a spring board for more experimental work, more systematic observations, and the more data that is gained, the more confident people can be about things like mechanisms.

Knowledge as a journey is central to Joe's EMK about science knowledge. He believes that science is not about the destination, but the process of creating and revising that knowledge. As such, communication plays an important role in his perception of how science is done. Each study serves as a springboard for more work, so science must be shared in order for it to flourish. Communicating this information leads to new data, contributing to increased confidence in the resulting mechanistic explanations of phenomena.

7.4.1. Instructor-guided hypothesis

Much like Connie, Joe uses information from the lab manual to formulate the instructor-guided hypothesis: "In both the lecture and lab we had learned about quantitative genetics and how a lot of genes." However, his reasoning for the hypothesis makes stronger connections between the data he presents and his claim that breeding parental plants with higher API will shift the offspring population toward higher API.

So our reasoning for that is because when we select for individuals with high API, and we cross those for the F2 select, we're driving the allele frequency up towards alleles that translate into deeper purple because of the additive effect. So we're using artificial selection. We're just taking the ones that have the alleles that translate to higher API. […] Because there's like 18 genes and there's different alleles that code for different shades of purple. The more purple individuals we're selecting for, the more alleles in all of those 18 genes that are going to be towards more purple, than on the less purple end, because we're not drawing from the pool of those alleles when we're selecting for the top 20 percent or whatever.

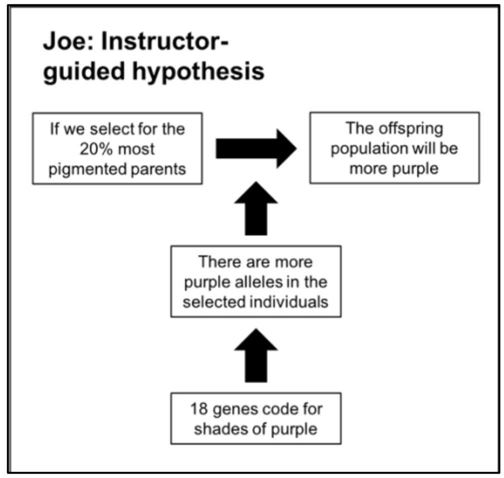

Figure 7.4. Joe's argument diagram for the instructor-guided hypothesis. Joe makes the claim that the offspring population will be more purple, given that the 20% most pigmented parents are bred to produce the offspring. He warrants this claim with the ideas that there are genes that code for shades of purple, and that the most purple plants contain more copies of these genes.

Joe's explanation makes a connection between choosing the most pigmented 20% parent plants and the generation of more pigmented offspring through the warrant that there will be more purple alleles in the selected plants compared to the random population. As we can see in Figure 7.4, all aspects of Joe's argument are connected in some way to the hypothesis.

7.4.2. Student-generated hypothesis

Joe shifts his information gathering from the lab manual to published literature as we saw with Connie above. Just as he does in the instructor-guided hypothesis, Joe connects the observation that anthocyanin stops radical chain reactions to his hypothesis that growing plants in BPA will make them more purple.

> *So we found a paper that, it was just a chemical study, no living organisms involved. It suggested that anthocyanin helps terminate radical chain reactions and these radical chain reactions can be caused by oxidative stress from BPA, as we found in another study. We then thought, okay, so we have this paper here that says that the anthocyanin somehow acts to stop damage from oxidative stress. So, could we perhaps see an increase in purple in the plants that are grown in the BPA environment? Because the BPA will cause an up regulation of the genes that are coding for anthocyanin production.*

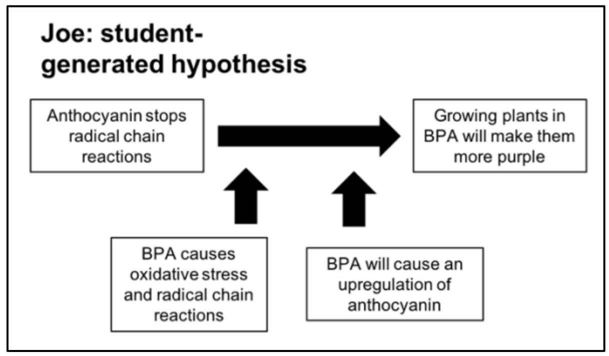

Figure 7.5. Joe's argument diagram for the student-generated hypothesis. Given that anthocyanin stops radical chain reactions, Joe claims that growing plants in BPA will make them more purple. He warrants this claim by stating that BPA causes radical chain reactions, and therefore will cause an upregulation of anthocyanin, a purple pigment.

Joe linked information he found in the literature to his hypothesis by warranting through other information he found in a paper. By linking BPA to oxidative stress and radical chain reactions, he was able to connect anthocyanin to BPA. Since Joe found these articles in papers, I asked why he looked at these sources for information.

Because the studies that were done, they were subject to peer review, first of all. If they were in a peer reviewed publication, then I feel more confident about the quality of the experiment that's described in the paper. That's the first thing. Second, if I see a study that has replicates, and it can be reproduced, the results can be reproduced, I feel better about the conclusions that are made from the study. Same thing as if it had more sample size, or stuff like that.

Joe responds by linking the validity of information to study replicates and reproducibility. Additionally, he feels that peer reviewed publications report higher quality experiments. This is an interesting contrast with Connie, who, when asked why she trusts peer-reviewed articles, explained that those were the instructors' expectations. Joe's perception of peer reviewed literature is linked to his EMK about how science knowledge: that scientists critique each other's studies to ensure that the results were reproducible, the study was appropriately designed, and the conclusions are appropriate.

I think it's important because that's the backbone of how scientific knowledge is generated. A big part of peer review is making sure that results are reproducible. That the conclusions that are drawn from a study are appropriate. That the study itself was designed in an appropriate manner and there's no conflicts of interest or something like that. If I see something that's peer reviewed, I can trust it more because I would think that somebody looked at the manuscript before it was published, looked at it to make sure it's actually that's something that results are reproducible, study was designed well, etc. Which I described earlier.

Joe trusts peer reviewed journals more because someone has checked over the manuscript to make sure that the results are valid. Someone also checks to ensure that the study was designed well, and that the conclusions drawn from the study were appropriate. As such, Joe's EMK about how science produces knowledge makes "peer review" a proxy for all the criteria that allow Joe to trust information from these articles.

A second warrant that Joe uses to connect anthocyanin to more purple colored plants is that BPA will cause an upregulation of anthocyanin. He explicitly states that this warrant is an inference.

> *The part of a mechanism in which my group, we inferred what was happening, was that we thought that the introduction of BPA would cause more anthocyanin production than if there was no BPA. That's the part we put in our hypothesis. That's the part we don't have anything to support, but given that we have studies that show that anthocyanin helps prevent damage from oxidative stress, and we have studies that show that BPA does introduce oxidative stress on plants, then we just put those together and made our hypothesis.*

Here, Joe is using an EMS, monitoring his certainty about knowledge. He is not certain that BPA causes upregulation of anthocyanin, saying that he has no evidence supporting the statement. However, he uses the information from papers to make a logical leap, which connects the information from the papers to his claim that plants exposed to BPA will be more purple in color.

7.4.3. Accepting or rejecting the hypotheses

Joe uses observations and significance to determine whether or not to accept his hypotheses. Joe calculates significance as instructed by the curriculum, by calculating the standard error, graphing his results, and checking to see if the error bars overlap. He accepts the instructor-guided hypothesis, but rejects the student-generated hypothesis.

> *What ended up happening is that in our final experiment, we found that it wasn't a statistically significant difference, but we did find that the BPA was affecting the plants, because when we took the secondary height data, we found that there was a significant difference in the height. The ones grown in the control environment were taller. […]*
>
> *We did find evidence to support our first hypothesis. So we concluded that the F2S population did have a higher API and we concluded that is because of allele frequency because of artificial selection.*

Consistent with all other participants, Joe evaluates only the data that was generated in the context of the classroom.

7.4.4. Joe's Conclusions

Although Joe rejected his student-generated hypothesis, he did notice that the plants grown with BPA were shorter than the plants grown without BPA. Joe infers that plant growth was inhibited by BPA based on this observation.

> *The ones grown in the control environment were taller. What can we infer from that? We can infer that the BPA was actually affecting the plants. So we know that our experimental treatment actually did something, anything at all.*

While Joe does not explain his interpretation any further, he does jump to a global perspective, contextualizing his conclusion within both science and society.

> *I think that when we saw that BPA actually negatively affected the plants growth, that's very important, because bringing it to the bigger picture, BPA is one of the most common environmental pollutants because it's in plastics so it will leach out into the environment and that's important because we want to know how, if that's going to negatively affect plants. Either agricultural crops, or natural preserves, or any kind of ecological restoration projects. Something like that. So I think that's significant.*

This global view of the implications of his results explains Joe's goal of attaining skills that facilitate communication of his results to the public. His conclusions led him to believe that BPA can affect agriculture, nature preserves, and restoration projects. As such, his experiments no longer pertain only to the classroom, but to society as well.

> *Then, I also think that even though we found no significant difference between those, the control and experimental groups, that's knowledge too. That tells us that right now there's no evidence that adding BPA is going to make plants more purple. That can be built on in future studies or studies done by other people. Just because there's no expected results, or any difference between the two doesn't mean it's not knowledge. It's all information that we can get from the study.*

Consistent with the outcome space, Joe's information gathering practices shift from classroom materials to published literature when he switches from the instructor-guided to student generated hypothesis. Whereas Connie's depended on input from instructors, Joe's practices relied more on his own reasoning. His EMK about science was also more prevalent in his discussions, and they influenced his reasoning for choosing to trust peer-reviewed information.

7.5. Composite 3: Transitioning

Participants applied Transitioning practices when they perceived differences between instructor-sourced information and information gathered through literature searches and experiments. Transitioning practices led participants to discuss dissonance between their classroom perspectives and their perspectives on how professional science constructs knowledge. Through these discussions, participants revealed metacognitive processes that helped to transition their argument practices from matching instructor expectations to constructing knowledge themselves. I use Cat as an exemplar for the Transitioning composite because she was open to trusting her own constructed knowledge over information she learned in the classroom. This openness facilitated her use of Transitioning practices in her accept/reject hypothesis and conclusion arguments.

7.6. Case 3 Cat: "It was interesting to discard the faith in what my teacher said based on my results"

Cat sits between Connie and Joe, at times deferring expertise to classroom experts and at other times trusting her own reasoning. Her discussions were a balance between the outcome of getting a good grade and the process of thinking through the research questions. For example, when asked about the importance of the project to her, Cat discussed both the quality of her project and the way she thought about the project.

> *I think the most important thing I got out of this proposal was learning that I didn't do enough work on it almost. Just working in a group element, my proposal that I actually submitted independently was so drastically different than my groups and it's just ... theirs had so much more information that was necessary for the project so it was interesting having to see the balance between aesthetic and information needed for the project.*

Cat focuses on the outcome of her project here, and similar to Connie, she assesses the quality of her project by comparing to other students' projects, as opposed to appealing to her own EMK about science knowledge as Joe does. Cat also discusses her thinking process as an important part of completing this project.

> *Just learning to like really think more in depth and like look at every aspect of a project from every possible perspective because you never know which perspective is the most effective of explaining the question that you're asking. There's also a lot of like looping around when thinking about a question. It's like you just like go through the same cycles but just with slight variation.*

Cat's description of her thinking belies a tolerance for ambiguity, to remain cognitively engaged with a subject for an extended period of time before settling on a conclusion. Her comfort with ambiguity is similar to Joe's, who believes that science knowledge is not an outcome, but a journey of revision.

7.6.1. Instructor-guided hypothesis

Cat juggles classroom and scientific contexts when she looks for information. For example, when looking at classroom content, she defers to expertise of the instructors.

> Interviewer*:* *So why do you trust the information from the lab manual?*
>
> Cat*:* *Because my grade was dependent on it. I feel like if they were going to grade my poster for accuracy, they wouldn't present unaccurate [sic] information as well… I would hope.*

While she also trusts the expertise of the authors of academic articles, she also pays attention to the methods presented and how they produced their results.

> Interviewer*:* *And why do you trust the academic articles?*
>
> Cat*:* *Because they just had very logical, well explained methods that produced results and I trust that people writing papers usually have pretty good understanding of what implications can be drawn from those results. I guess it's just trust and experience.*

Cat felt that the instructor-guided hypothesis was a control that helped support the validity of the student-generated hypotheses. "I mean, it was entirely designed that way, that we had like a genetic marker type thing. So we could trust our like [student-generated hypothesis] results more." As such, Cat approached the construction of the instructor-guided hypothesis as an assignment with a predetermined answer.

> *We talked about it a lot in discussion, like it was one of those roundabout conversations where we were really looking for specific keywords in talking about this. It makes sense logically because the actual allelic variations are reduced to an extent to only have the like more purple alleles within the population but since it is a quantitative trait you can't know that for sure, that all of the allelic variation for the green or like lighter colored API aren't still in the population of the new seeds if that makes sense.*

Here again, we see Cat switching between classroom and scientific perspectives. In the beginning of the above excerpt, she speaks as if this discussion was merely an exercise that she needed to complete. However, she begins to reason through the validity of the hypothesis, and comments on the uncertainty of genetic variation.

Cat explains the reasoning behind the instructor-guided hypothesis with little extraneous information. She connects the idea that selecting the 20% most pigmented plants as parents will condense the genetic variation, making it more likely that there will be a higher number of pigment alleles in the offspring.

> *Yeah, so we talked a lot about like the graph of what the parental generation looked like in terms of its API and how if you like slice off the top 20% and how his has a more condensed group of allelic variation […] The alleles, they just... they're additive so one allele might add like two units of purple and another allele might add like four units of purple to the API, units being arbitrary, I don't know what*

units they are. The alleles, if you have a lot more alleles that are adding more purple, it's going to have a higher API.

Cat attributes the information in this reasoning to discussions she had in class. Additionally, Cat reflects upon selecting smaller fractions of the parental generation.

…but you still have a lot of different alleles that will code for a lighter color of API and we just talked about oh, how does that compare to a 1% sliver or even a 5% sliver type thing. And just talking about the different types of crosses that we could arise at because it's a quantitative trait and there is a lot of different variables obviously.

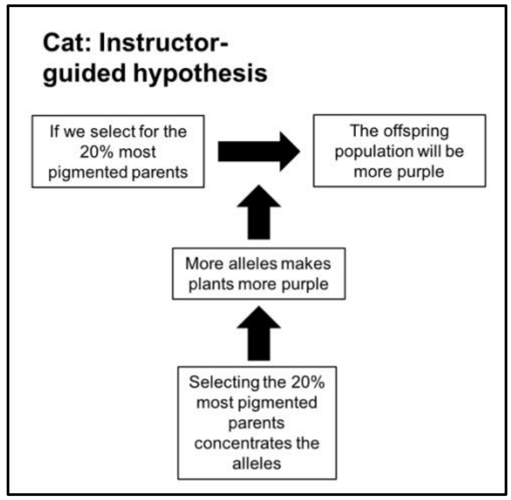

Figure 7.6. Cat's argument diagram for the instructor-guided hypothesis. Cat's data and claim are identical to Joe's, although her warrants are slightly different.

This discussion is an example of how Cat keeps herself in an open state as she thinks about the project. Even after establishing the hypothesis, she continues to reflect on other possibilities, as opposed to ceasing her reflection once an answer has been established.

7.6.2. Student-generated hypothesis

Consistent with the outcome space and the other cases presented here, Cat transitioned to using academic journal articles for the student-generated hypothesis.

We just saw that other different types of stressors acting on Brassica rapa or even monoculture plants would just induce stress and so we said oh, maybe root density *is enough of a stressor to also induce a higher API as shown in other studies of Brassica plants. Yeah, a lot of our resources talked about the effects of nutrient competition, water competition, and light competition but we didn't really find any articles that talked a whole lot about root density competition so we walked in with that root density was an assumption more than as part of our biological rationale.*

As summarized in Figure 7.7, Cat connects data suggesting that nutrient, water, and light competition increase anthocyanin production in *B.rapa*. Additionally, she presents research on her poster supporting that these types of competition cause stress in plants. As part of her reasoning, Cat explicitly states that while she believes that root density competition would cause stress, she has no evidence saying so. This is an example of Cat employing an EMS, specifically, monitoring her level of certainty. Since she is not certain about root competition causing stress in *B. rapa*, it is not certain that tightly packing plants in a bottle will result in higher pigmentation. Therefore, this hypothesis is worth investigating.

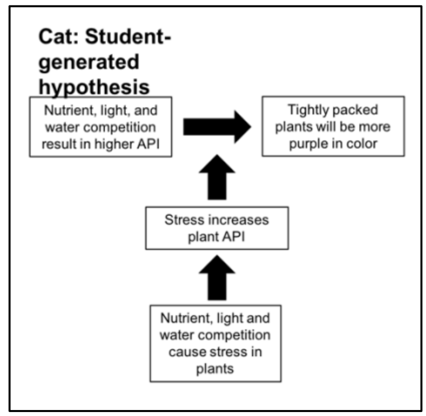

Figure 7.7. Cat's argument for the student-generated hypothesis. Given that resource competition results in higher API, Cat claims that tightly packing plants will result in plants that are more purple. She warrants this claim by stating that resource competition causes stress in plants, which, in turn, increases API.

7.6.3. Accepting or rejecting the hypotheses

Deciding whether to accept or reject her hypotheses was the part of the research cycle where Cat reconciled the differences between her classroom and scientific perspectives. She expected to accept the instructor-guided hypothesis, but instead rejected it.

> *It was just entirely like didn't support our hypothesis. It supported alternative for both [hypotheses], which I was very surprised about [the instructor-guided hypothesis] because I had definitely thought that our hypothesis would be dead on and that 20% selection would be enough but it wasn't enough to show significant differences in API.*

Cat decided to reject both hypotheses based on the same calculations Joe made in his project. She calculated the standard error, graphed her results, and checked for overlap in her error bars between the populations.

> *What tells me that it's not sufficient is that there wasn't a significant mean difference between the API ratings of the selected versus randomly selected seedlings. We used specific error bars and I compared the graphs and we saw that error bars overlapped so we couldn't say that there was a significant difference.*

Rejecting the hypothesis led Cat to question the certainty of what her instructors were saying in class.

> *Just looking at the graph, 20% seemed like such a small amount but it just wasn't small enough. That's pretty funny. And just the way that the professors were presenting it almost made it seem like that is what we should have gotten. It was very interesting to like almost discard the faith in what the teacher said just based on my results and it's like oh, were my results really accurate with such a small sample size or should I have gotten a difference in the phenotype?*

She also questioned her own results, reasoning that her sample size may have been too small. This is an example of EMS, where Cat is *evaluating* the validity of her results. So she checked her results with her classmates, who had the same results as she did.

> *I think I had to trust the science because other groups also got those results so adding to the sample size increased confidence. I'd say yeah, it's nice to like see it with my own eyes what the teachers are saying versus what they're not saying.*

Cat's decision is an interesting shift from seeing the instructor as the source of knowledge in the classroom to seeing her own experimental data as a source for knowledge. However, Cat does not consider the knowledge she produced as certain. She recognizes that while alignment between her results and another group's results doubled her sample size and increased her confidence in her data, her confidence would be higher if she compared her results to the results of the entire class. Here again, Cat applies the EMS of evaluating the validity of her data, and applies her EMK that, she can be more confident about her experimental data if her sample size is larger.

> Interviewer*: So to be perfectly clear, when you increased the sample size and looked at the whole classroom, was the hypothesis still rejected or was it accepted? Do you know?*
>
> Cat*: I actually don't know. I just know that one other group I compared with, they had the same results as we did so that doubled my own group sample size ... So it doubled my confidence. But it didn't increase my confidence to the size of my class entirely.*

Cat uses her understanding of error and sample size to explain why she feels uncertain about her results.

> *Because there's less variation that's like ... I suppose an extreme case that could be found like one plant that should have been average ended up being like very, very light purple, that little bit of a skew is going to drag the mean just so drastically with a small population as compared to a bigger population.*

She explains that an extreme case in a smaller sample size can skew her results. As such, increasing the sample size helps increase her confidence in her results. Unfortunately, she only checked her results with one other group, which did increase her confidence, but not much.

7.6.4. Cat's Conclusions

Cat rejected her student-generated hypothesis and reasoned that root density was not enough of a stressor to affect anthocyanin production in her plants. This was the alternative hypothesis that she presented in her poster.

> *The conclusion was that the root density did not ... it was not enough of a stressor to actually change the average anthocyanin production of our plants. It's also possible that we just didn't compact them enough since they are very small plants and they are specifically bred to grow and thrive in tiny conditions. One of the professors said that he tried growing a whole bunch of* Brassica rapa *in like a little film bottle and he could fit about 30 of them in there and they weren't doing well but they survived, so it's possible that we didn't actually cram them in enough. We*

thought that cramming them in four times the amount of the control should have had some impact at least but we didn't find that.

It is also interesting to see that Cat is open to other explanations for her results. In the previous section, Cat seems to accept the alternative hypothesis simply because she rejected the expected hypothesis. In the above excerpt, Cat proposes another explanation, that the plants may not have been packed closely enough to even induce stress. She cites information from a professor who was able to grow 30 plants in a container half the size Cat used (she tried to grow 24 plants in a container). Cat is open to this new information, and is willing to change her conclusions to accommodate the new evidence.

7.7. Cross-case and composite analysis

The three cases presented here differed in the ways that they viewed the certainty of knowledge, themselves as knowledge sources and EMK about science knowledge. Connie sees others as knowledge sources, so she seeks information from her instructors and peers. She trusts this knowledge so strongly that she will report the information even with an incomplete understanding of the reasoning behind the claims. We see examples of this trust in her explanations of the instructor-guided and student generated hypothesis, as well as her discussion of etiolation. Connie's EMK about scientific knowledge seems to be centered around an expert who imparts the knowledge to her. She matches this EMK in her descriptions of the hypotheses, listing scientific facts with little explanation of how they are connected to each other or the claim.

However, when Connie decides whether to accept or reject her hypotheses, she relies on her own interpretations to make the decisions. She discusses these results in less certain terms, using phrases like "I think…" before many of her claims. When asked what part of her project she considers knowledge, Connie points to the results of her experiments. She feels that she can pass that knowledge onto someone else, therefore it is knowledge.

> *I mean, if they wanted to know the difference of Anthocyanin pigment intensity, they could look at the results and see. So, I think that would be like, someone could gain knowledge from that. And then a conclusion kind of basically explains the results that we have.*

The above excerpt matches well with Connie's EMK that knowledge is propagated, passed from one person to the next. In this case, Connie is able to pass what she learned in her experiment to others, therefore she believes it is knowledge. In much of the research cycle, this EMK manifests from a classroom perspective, where knowledge is passed from instructor to student. The context is different when Connie is analyzing her own experimental data. While she may or may not realize it, Connie is the one constructing knowledge when she performs these experiments. Having an explicit discussion about knowledge construction at this stage of the research cycle may help broaden Connie's EMK about the construction of science knowledge, and who can construct this knowledge. These discussions may help Connie broaden her practices to include Transitioning and Enculturing practices as well.

Connie's perception of knowledge is at once similar and different from Joe's EMK about science knowledge. While both believe that knowledge is passed from one individual to another, their perceptions of the certainty of knowledge make these perceptions very different. Connie passes knowledge on as a completed outcome, but Joe passes knowledge to contribute to an ever-evolving entity. We see him using uncertainty to explain the reasoning behind his hypotheses. Joe explicitly points out where he is making inferences, which he connects to the generation of his hypotheses.

However, it seems that Joe feels there are varying levels of certainty in science knowledge. For example, when searching for information, Joe trusts information from published literature. Again, on the surface this looks similar to Connie's practice, but their reasons for searching in published literature databases is where these practices diverge. Connie states that she looks in these databases because authorities in the classroom instructed her to do so, while Joe feels that peer-review forms the backbone of how science knowledge is generated.

Another notable difference between Joe and Connie's discussion was that Joe brought his conclusions out of the classroom context into a societal context. He felt that he had constructed knowledge in the course of the project because he had revealed information that could impact society: BPA is a common pollutant that he now understood could affect the growth of plants. While Connie kept her conclusions in the context of the classroom, staying with the instructor suggestion that her data were a result of etiolation, Joe looked into other contexts, seeing the importance of his work with respect to agriculture and ecological restoration.

Cat viewed her epistemic practices with a mixture between the classroom and scientific contexts. She trusts the instructors because they give her the grades, yet she also feels that thinking through a question from many different perspectives is an important part of constructing knowledge. Cat keeps these perspectives separate while constructing her two hypotheses, however they clash when she analyzes her data, seeing information from her instructors conflicting with what her data are telling her. Cat's comments make it seem like she felt knowledge from the instructors was certain, however rejecting the instructor-guided hypothesis made her question that certainty. Cat was open to the uncertainty that came with rejecting the instructor-guided hypothesis.

Cat also felt that learning something made it knowledge. When asked what part her project she considered knowledge, she talked about what she learned from the project.

> *I learned a lot about needing to know the actual biological mechanisms that occur behind these differences and I learned a lot more about what it means to really understand what's going on in a system.*

Asked about what understanding means to her, Cat reconnects with the idea she discussed at the beginning of the interview: the importance of viewing a question from many different perspectives.

> *What it means to understand something is be able to look at a system or a question from multiple perspectives and to finally reach a conclusion that's like a culmination of all these different ideas in a way that's like logical and actually representative of what's going on I suppose. With the Brassica Rapa experiment in particular, you can look at it from the biological mechanism where it's talking*

about like genes producing, but you can also look at it from environmental impacts where it's like the stressors are going to increase anthocyanin production. You can also look at it from like the allelic variation so you have to look at all the different perspectives to really understand something and gain knowledge about it.

This definition of knowledge is an interesting departure from her earlier comments about trusting instructors because they held sway over Cat's grades. Like Connie, Cat's most salient departure from a classroom perspective occurred while she was deciding whether or not to accept her hypotheses, reinforcing the importance of authentic inquiry in development of students' epistemic practices.

7.8. Relationship between composites

Describing each composite and exemplar case individually revealed relationships between Validating, Transitioning, and Enculturing practices. If the classroom knowledge culture is viewed as a community of practice, Validating practices can be seen as a way for participants to move from outside into the community. Transitioning practices help participants to engage with metacognitive processes that help to develop EMK about how the community constructs knowledge. This helps participants to move toward the center of the community of practice. Finally, Enculturing practices are used by participants that understand the knowledge processes of the community. Since individuals using Enculturing practices were at one time outside the community of practice, they would have used, or may still use Validating and Transitioning practices in this community of practice.

Therefore, Validating and Transitioning practices are encompassed by Enculturing practices, as depicted in Figure 7.8.

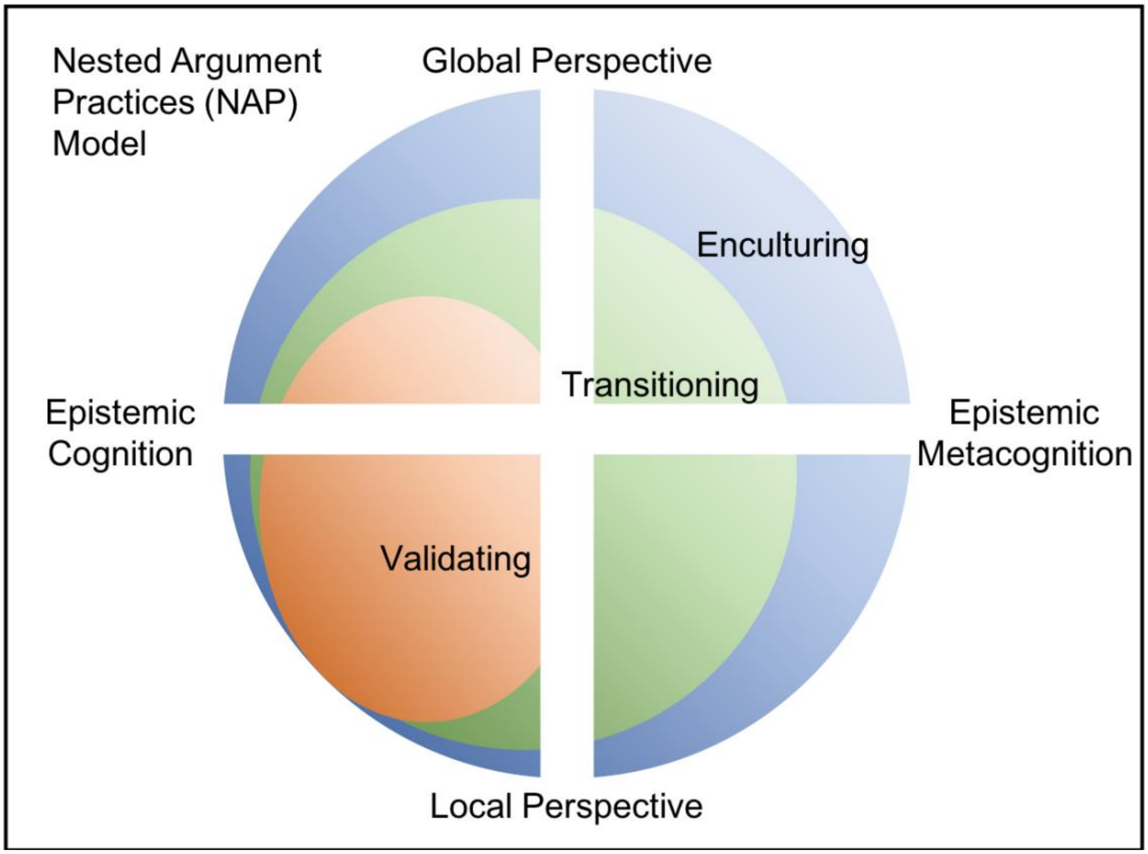

Figure 7.8. The Nested Argumentation Practices (NAP) Model. Individuals wanting to become part of a community of practice enter the community using Validating practices. Transitioning practices help individuals become more central to the community, and central participants use Enculturing practices. As depicted in this figure, Validating practices reside in the Epistemic Cognition portion of the Epistemic Practices (EP) framework. Transitioning practices move into the Epistemic Cognition portion, helping individuals develop EMK about knowledge construction in the community of practice, facilitating the use of Enculturing practices.

7.9. Hypothesized Composite: Switching

During analysis of the three composites described above, a fourth, hypothesized, composite emerged from one participant, Hannah, that did not fit well with any of the other three composites. While some of Hanna's practices fit with the Validating and Enculturing composites, her practices were not nested within the progression of the other three composites. That is, while Hannah aligned her beliefs about scientific knowledge production with her perception of classroom expectations, she did so in a way that allowed her to maintain her antecedent religious beliefs without fully committing to scientific knowledge culture.

> *So like a theory of evolution is based on a vast, vast, vast sum of knowledge, which is like the experiments that people have done and the observations that they have made. But the theory of evolution itself isn't… like, it's not, we can't ever know if it's correct or not. […] Which I guess in my mind prevents you from, prevents me from classifying it as knowledge. But in practice because it's been supported so many, so many, so many times, you can like when you're doing other science or when you're just like living your everyday scientific life, you could think about it as correct. Because there's nothing that we've seen that has disproven it. But because it can't be proven correct, in my mind I guess you can't say that its' knowledge. Yeah that's how I feel about it.*

In the excerpt above, Hannah describes scientific theory as something based on a sum of experiments and observations scientists have made. However, she keeps this definition at arm's length by focusing on the inherent uncertainty of scientific theory. She feels that in a scientific context, the theory can be thought of as "correct," but because theories have uncertainty, they cannot be considered knowledge. Hannah explains that her view of scientific theories is influenced by her religious beliefs.

I think that it might be important to note that there are probably beliefs that I have outside the scope of this course that effect like how I think about the reliability of scientific data and like theories I guess, so foundational science. Yeah I don't know I guess that's probably, that might be important. So I talked about evolution, so I'm a Christian and therefore like those beliefs that I have might cloud my judgment as to whether I consider the theory of evolution knowledge. To someone else who doesn't have that belief then maybe the theory of evolution is knowledge.

The distinction between Hannah’s perception of what scientists consider knowledge and what she considers knowledge is what separates Hannah from the other three composites. She searched for peer reviewed articles because “…Well because first of all, we were told we had to,” but also because of her perception of how scientific knowledge is constructed.

So a bunch of people did a bunch of that over and over again, which is science as we know it. And they published their papers, their results in journals and the process by which that occurs is that they do to experiment, they write the paper, then they send it to the people who publish the journal and those people look at it and read it. And they're also scientists and they say, well this seems reasonable but did you ever think about like for instance this other factor that maybe they hadn't considered. […] And so that process of scientist to scientist feedback helps to ensure that what is actually published in the journal is reasonable and is valid. So that when us as like students or scientists later on go back to look at it, we can find the results, see the conclusions that they draw and be reasonably certain that that information is factual.

It is clear from the above excerpt that Hannah has a clear understanding of how science constructs knowledge. Therefore, Hannah's practice of searching for peer-reviewed literature could be considered on the surface as a Validating (meeting instructor expectations) and an Enculturing (doing what scientists do) practice. However, since Hannah does not believe that what is produced from these practices is knowledge, she is

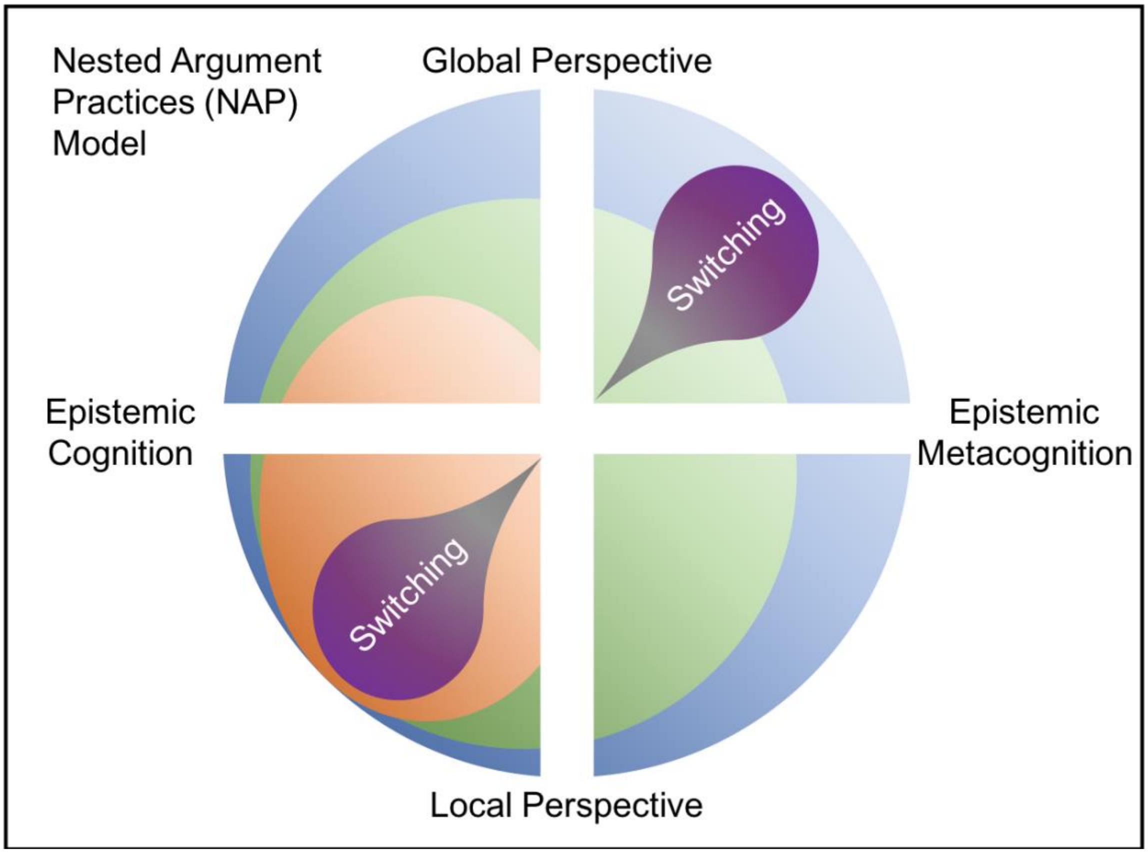

Figure 7.9. The NAP Model with hypothesized fourth composite. Hannah applies practices in the Switching composite. She uses practices in the epistemic cognition/local quadrant to match instructor expectations, and practices in the epistemic metacognition/global quadrant to reconcile her EMK about how scientists (not Hanna) construct knowledge with her religious beliefs.

not committing to the knowledge culture of the classroom. Instead, I hypothesize that Hannah switches between two separate contexts (Figure 7.9). She uses Validating practices to match her perceptions of instructor expectations so that she can progress in the science curriculum, but also uses Enculturing practices to reconcile what she is learning in science curricula and her religious beliefs. As a result, Hannah is able to advance in her scientific coursework while maintaining her religious belief that the theory of evolution is untrue.

This hypothesized Switching composite provides insight into the interaction between science education and religious belief. It also reveals a limitation of my study: religious beliefs were not one of the variables I considered as part of my maximum variation selection criteria. However, the Switching composite does present an avenue for future research into how students negotiate classroom culture and the cultures they bring into the classroom.

8. DISCUSSION

In this chapter, I situate my results in the field of biology education by first describing how participants' descriptions of their epistemic cognitive and metacognitive practices fit with contemporary biology instructional practice. Following this discussion, I describe how biology competencies laid out by AAAS are reflected in the analytic framework and outcome space. The connections made between my results and contemporary biology teaching and learning demonstrate the importance of considering epistemic cognition and metacognition in general biology curricula. To facilitate the inclusion of these important subjects, I suggest some strategies to leverage results of this study for instructional practice. As part of this discussion, I describe how student consideration of epistemic cognition and metacognition can result in adaptive expertise, the ability to apply procedural skills across contexts (Hatano & Inagaki, 1986). Finally, I discuss implications of this study for biology education research by considering connections to other theoretical constructs

8.1. Cognitive Outcomes and Metacognitive Processes

Undergraduate research experiences such as CUREs reflect biology education's shift towards teaching students *how to do science* as opposed to *what to know about science* (Duschl, 2008). However, focusing solely on the actions of science may result in rote performance of practices with limited understanding of how science constructs knowledge (Berland et al., 2016). One challenge for assessing understanding of science knowledge through practices is that students may use similar practices for diverse epistemic reasons. For example, all participants in my study searched for peer-reviewed literature to construct their student-generated hypotheses, but while some students felt the information in these articles was credible because its validity was reviewed by experts in the field, other students felt the information was reliable because searching for information in peer-reviewed articles was required by the instructors. These results parallel work done in K-12 education, where researchers observed high school genetics students constructing arguments to either satisfy the constraints of the assignment (doing the lesson), or to satisfy scientific conventions of valid knowledge (doing science) (Jiménez-Aleixandre, Bugallo Rodrìguez, & Duschl, 2000).

An example of the tension between "doing the lesson" and "doing science" manifests in Chapter 6, with Bill's decision to ultimately accept and reject the instructor-guided and student-generated hypotheses, respectively. Bill rejects the instructor-guided hypothesis based on his experimental data, but still believes that the hypothesis is true because of the genetics concepts he learned in class. His classroom knowledge supplies a correct answer, an *outcome* that he can work towards while "doing a lesson." On the other hand, Bill rejects his student-generated hypothesis, citing uncertainty in his experimental and information gathering *process* while "doing science." As discussed in Chapter 6, Bill seems to epistemically frame the instructor-guided hypothesis as certain, and the student-guided hypothesis as uncertain. The correlation between "doing the lesson" and certainty, and "doing science" and uncertainty makes sense, as lessons tend to encourage correct answers, which are inherently certain, and science seeks to describe natural phenomena, which are often unknown and therefore uncertain.

The differences in Bill's epistemic framing of the two hypotheses is accompanied by differences in the way Bill discusses his epistemic practices. While discussing the instructor-guided hypothesis, Bill **validates his conclusion** with information from an expert. Since his conclusion was the object of the statement, he used epistemic cognition to pursue the *outcome* of validating information. On the other hand, Bill discusses **his own certainty** about his experimental plan and his background knowledge when talking about the student-generated hypothesis. Since his own certainty about the experimental plan and knowledge was the object of the statement, Bill was using epistemic metacognition to discuss his experimental *process*. The observed connections between epistemic cognition and outcomes, and epistemic metacognition and scientific process are important, especially if we re-visit the shift in biology education from teaching content knowledge to scientific processes. If we truly want students to consider the *process* of scientific knowledge production in addition to *outcomes* such as valid conclusions drawn by experts, we should encourage students to bear in mind both epistemic cognition and metacognition.

8.2. Competencies for Biology Literacy are Reflected in the Outcome Space

"Doing the lesson" is not an inherently unproductive practice, especially if the lesson aligns with the knowledge practices of the discipline. Accordingly, biology educators have set forth several core competencies. Those discussed by participants in this study include the ability to apply the process of science, reason quantitatively, tap into the interdisciplinary nature of science, communicate and collaborate with other disciplines, and understand the relationship between science and society (Brewer & Smith, 2011). Participants in this study discussed each of these competencies in a diversity of ways.

8.2.1. Applying the Process of Science

According to the AAAS Vision and Change in Undergraduate Biology Education, "[s]tudying biology means practicing the skills of posing problems, generating hypotheses, designing experiments, observing nature, testing hypotheses, interpreting and evaluating data, and determining how to follow up on the findings" (AAAS, 2011, p. 14). All participants applied the process of science according to the research cycle as prescribed by the instructors. However, the outcome space exposes differences in how participants described these skills. Students in the local quadrants of the outcome space used phrases such as "they told us to…" or "we are supposed to…," engaging the process of science by "doing the lesson." However, participants in the global/EMC quadrants of the outcome space engaged the process of science by "doing science," generating hypotheses, designing experiments, and analyzing data in particular ways because they fit with their perspective of how science is done.

8.2.2. Reasoning Quantitatively

As shown in the outcome space, students did not discuss their experimental data in terms of anything outside of classroom-based practices, therefore they did not discuss quantitative reasoning in the global quadrants either. Students who discussed quantitative reasoning did so by talking about using chi-squared tests to check whether or not hereditary patterns matched Mendelian genetics, and about using standard error to determine if there was a significant difference between experimental means. Student discussions in the EC quadrant focused on the acceptability of their hypotheses. However, students who discussed quantitative reasoning in the EMC quadrant considered their own certainty about drawing conclusions from their data. For example, Cat explained her thinking about how overlapping error bars affect her confidence in her conclusions.

> Overlap of the error bars means that the difference isn't so drastic that you can draw a conclusion from it saying that it is different or it means that if you do try and draw a conclusion, you're not going to be very confident in it, in saying oh, this is definitely because of this factor I'm proposing but instead it could have happened randomly.

Cat's discussion above demonstrates her engagement with one of the aspects of the Nature of Science, that scientific knowledge varies in its certainty (Norman G Lederman, 2007; W. A. Sandoval, 2005). It is important to note that students who did not discuss quantitative reasoning in the EMC quadrant may have also considered the certainty of their conclusions. However, I cannot make concrete judgements about participant quantitative reasoning without EMC discussions.

8.2.3. Tapping into the Interdisciplinary Nature of Science

Many biological innovations stemmed from integrating knowledge between biology subdisciplines, and between biology and other disciplines. When students discussed tapping into the interdisciplinary nature of science, they examined it in terms of integrating published information from different disciplines to construct or revise their hypotheses or conclusions. These student discussions of interdisciplinary science appeared in the global/EMC quadrant of the outcome space. Alyssa provides an example of discussion about using interdisciplinary knowledge in revising her conclusions. In the initial construction of her hypothesis, Alyssa chose publications that provided information she perceived as closely related to her topic of interest.

> I looked at how closely related it was to what I was looking at. In some of them, it was looking at a pigment that had been isolated from a plant. I didn't think that would be as helpful, because we're looking at it still in the production process.

Alyssa found some information from papers that studied isolated pigment rather than the whole plant, and decided that the information was not relevant to her project. Following data collection and rejection of her student generated hypothesis, Alyssa revisited the literature that reported on the isolated pigment, this time incorporating it into her reasoning.

> Because the reason that it synthesizes the pigment is different between the *Brassica rapa* and the apples. I don't think I did enough research into it, because I think it's a different mechanism that causes it. […] Yeah, because it was mostly like the isolated ... They looked more at the isolated pigment once it's been extracted from the plant and how that reacts to temperature. It was more into the chemistry of it.

Alyssa uses her EMS to monitor changes in her understanding here ("I don't think I did enough research into it") to realize that her initial judgement of the isolated pigment study was incorrect. This reflection helps her recognize that relevant information includes sources from the chemistry discipline.

8.2.4. Communicating and Collaborating with Other Disciplines

Other participants discussed interdisciplinary knowledge in terms of posing new problems or constructing new knowledge. Dan felt it was important to connect knowledge from different (sub)disciplines, which he called “bases of knowledge.”

> I think the importance of connecting different bases of knowledge would be to see if you could ask questions that might be able to connect knowledge that wasn't previously known. By combining connections between all this actually did help create more knowledge and lead to more questions. The connection's a way to help generate more knowledge.

Connecting knowledge bases is part of Dan’s perception of how knowledge is constructed in biology, in other words, his EMK about biology epistemology. As such, his discourse about knowledge construction and knowledge from outside the classroom context place this discussion in the global/EMC quadrant of the outcome space.

8.2.5. Understanding the Relationship Between Science and Society

Connecting knowledge bases is an essential part of understanding the relationship between science and society. We saw Joe engage with this competency in Chapter 7 when he contextualized his results with agriculture and ecological restoration. Joe discussed the significance of his data in terms of how it would affect society. Additionally, he felt that this connection to society was one reason he considered the results and conclusions from his project to be knowledge. Therefore, Joe's discussions on this topic fall within the global/EMC quadrant of the outcome space.

Examining the outcome space through the lens of the biology literacy competencies reveals the importance of the global and EMC quadrants. When students described their engagement with the competencies of tapping into the interdisciplinary nature of science, communicating and collaborating with other disciplines, and understanding the relationship between science and society, their discussions fell into the global/EMC quadrant of the outcome space. Furthermore, while participant discussions about quantitative reasoning were located in the local/EC quadrant, students should ideally consider how their quantitative analyses interact with contexts outside the classroom, and how to evaluate published quantitative analyses. The connections between participants' global/EMC discussions and the biological competencies underline the importance of encouraging students to consider perspectives beyond the classroom in epistemic metacognitive ways.

8.3. Implications for Teaching Practice

8.3.1. Balancing Effortful Metacognition with Timely Learning

This is not to say that students should think globally and epistemically at all times. Thinking metacognitively about knowledge takes effort and time, and, in some contexts, may not be as productive as taking cognitive shortcuts such as trusting an authority. For example, when learning about biology theory in a classroom, students should trust instructor expertise instead of questioning its validity each time students engage a new theory. Indeed, research in nursing decision-making paint cognitive shortcuts in a positive light as they reduce the cognitive load, allowing nurses to make better decisions quickly (Ferrario, 2004). However, using cognitive shortcuts in scientific inquiry can be problematic if they are not grounded in scientific practice. For example, one practice Connie used to validate information was to check the format of the source material: "…usually it was formatted as how we had our proposal formatted or our papers with the abstract and all that." Used alone, checking formatting as a cognitive shortcut would be an unreliable practice to validate scientific information.

Thus, teaching students to use cognitive shortcuts or heuristics grounded in science epistemology should be a priority when presenting students with the process of science. However, results of this study suggest that monitoring students' epistemic practices may not accurately represent their understanding, since students who used similar practices did so for both scientific and non-scientific reasons. In this study, evidence that students achieved the competencies for biology literacy only appeared in conversations that fell into the global or EMC quadrants of the outcome space. Therefore, to ensure that students' practices are grounded in science epistemology, instructors should have explicit classroom discussions about the biology epistemology behind epistemic practices. The analytic framework provides guidance to instructors for opportunities to engage students in these kinds of discussions through the description of specific argument contexts in which students discussed their practices globally and with epistemic metacognition.

8.3.2. Leveraging the Outcome Space for Instructional Practice

If the goal of instructors is to educate students about the global and epistemic metacognitive aspects of biology practice, it makes sense to find contexts in which students are already engaging with these types of discussions. The outcome space indicates that students most often engaged in global/EMC discussions while talking about their student-generated hypotheses and conclusions. As participants shifted their information gathering from course materials to published peer-reviewed articles, they contemplated the credibility of information from these new sources. While some participants such as Hannah stated they used peer-reviewed articles because it was a requirement of the assignment, others like Joe and Nick discussed the importance of peer-review to knowledge construction in science. This provides an opportunity for instructors to connect the classroom practice of requiring peer-reviewed sources to the scientific practice of socially constructing knowledge (Longino, 2002) through peer-review. Taking steps like these to maintain the balance between classroom and scientific practice is essential for meaningful science teaching and learning (Berland et al., 2016).

Students also argued their conclusions in global and epistemic metacognitive ways. For example, in explaining why her hypothesis was not accepted, Alyssa finds that information she initially determined as irrelevant became an essential part of her conclusion. Likewise, as Cat considered the reasoning behind her results, she reflected that information she learned in class about heredity was not as certain as she initially believed. Situations where students build arguments like these are an opportunity for instructors to encourage discussion of the importance of iteration in scientific knowledge construction.

Additionally, when students perform experiments in class, they may encounter results that do not match established biological theory due to the stochastic nature of biological processes. The randomness of nature gives instructors an opportunity to discuss with students the varying degrees of certainty in biological knowledge.

Participants discussed certainty of knowledge in both student-generated hypothesis and conclusion arguments. Uncertainty about claims drove participants like Jennifer, Bill and Joe towards hypothesis formation. They identified knowledge gaps where there was little published evidence to support their claims, and formed hypotheses to explain the outcome of their predicted results. At this point, instructors asked students to construct a BR to explain their hypotheses by making logical connections between published literature and students' assumptions about the biological systems they were working with. In constructing their BRs, participants explicitly identified inferences, assumptions and knowledge gaps, facilitating their engagement with both certain and uncertain knowledge. Students returned to these BRs when it was time for them to interpret data and draw conclusions from their results. Participants who accepted their hypotheses used the BR to explain their results, while those that rejected their hypotheses used the BR as a starting point to construct explanations for their unexpected results. The BR formed a scaffold for students to reason through their hypotheses and conclusions.

The BR provides a means for instructors to assess and give feedback on student engagement with biology epistemology, something that instructors in this curriculum leveraged through low-stakes ungraded feedback presentations. As students presented their BRs in these feedback sessions, instructors asked questions such as "How relevant are results from model systems other than *Brassica rapa* to your studies?" to model biological reasoning processes. Asking these kinds of questions encourage students to engage in reflexivity, an internal conversation whereby an individual evaluates and re-evaluates their actions and decision making (Archer, 2012). Engaging students in conversations such as those in feedback presentations teaches students the contextual processes required to use informed reflexivity to arrive at favorable epistemic ends such as knowledge and understanding (Weinstock, Kienhues, Feucht, & Ryan, 2017).

8.3.3. Fostering Adaptive Expertise through Epistemic Discussions

Reflexivity sets students on a path toward adaptive expertise, wherein individuals are able to address novel and complex situations by solving problems through the generation of new knowledge (Mylopoulos & Woods, 2017). Hatano and Inagaki (1986) propose that there are two kinds of expertise: routine expertise, and adaptive expertise. Individuals with routine expertise have vast procedural skill, and are able to perform a task with automaticity as long as the object and constraints remain consistent. Individuals with adaptive expertise, on the other hand, combine conceptual knowledge with procedural skill, and are therefore able to apply their procedural skills in different contexts (Hatano & Inagaki, 1986). The ability to apply procedural skills across contexts is an important outcome for biology students, especially if they are expected to be able to tap into the interdisciplinary nature of science, collaborate across disciplines, and connect science to society (Brewer & Smith, 2011). Furthermore, biology students aiming to transition into professional disciplines such as medicine will need the ability to efficiently solve problems by integrating both scientific knowledge and social knowledge such as patient experience (Mylopoulos, Kulasegaram, & Woods, 2018).

How then, are we to foster adaptive expertise, and how can discussions about science epistemology get us there? To answer this question, we return to Hatano and Inagaki (1986), who propose three factors contributing to the development of adaptive expertise: variation in parameters, context, and understanding of the system. Exposing an individual to a variety of parameters motivates them to modify their skills to fit different contexts (Hatano & Inagaki, 1986). In this study, students were exposed to genetic experimentation, a system that inherently contains randomness. Some students, such as Cat, recognized the variation between her experimental observations and what she learned about genetics in class. Through a reflexive process, Cat considered her own certainty in classroom knowledge alongside the experimental results, which, in turn, modified her thinking about the certainty of classroom knowledge. This instance provides an opportunity to enhance Cat's procedural and epistemic skill of trusting scientific experts by discussing how to determine when it is appropriate to question or trust a scientific expert.

The second factor, context, refers to the epistemic (or non-epistemic) goals of applying a procedural skill in a particular context. When performing procedural skills for a reward, such as a good grade, people are more likely to choose skills that are safer or more algorithmic (Hatano & Inagaki, 1986), making it less likely for them to attempt a diverse set of skills that may be applicable across contexts. This was the case with Connie, who based her research actions on what instructors told her to do. Additionally, Connie had an opportunity to modify her procedural skills when she saw unexpected results, but opted to once again interpret her results according to the instructor's advice. Connie's actions reflect the contexts described in the Epistemologies in Practice (EIP) framework, where practices can be seen as meaningful on a continuum between classroom and scientific contexts (Berland et al., 2016). Shifting the goal from classroom-based performance toward scientific knowledge construction can motivate students to experiment with multiple approaches, preparing them to manage variable contexts when they encounter them.

Understanding the system behind the procedural skill is the final factor contributing to the development of adaptive expertise. When working with the goal of understanding, individuals engage in mental experimentation, integrating ideas that may not yet be fully formed to explain phenomena. For example, Joe searches for peer-reviewed information (procedural skill) with an understanding of science epistemology (system). He understands that the information he finds in published literature is reliable because of his EMK about peer review, and constructs his hypotheses based on assumptions and inferences he makes from the information he finds. Understanding the system behind peer-review helps students to determine the reliability of information so they can make informed decisions about science.

While there are examples in the data of students interacting with factors that lead to the development of adaptive expertise, the data do not show what, if any, aspects of the CURE encouraged students to engage with these factors, or what kinds of discourse encouraged students to consider their project through a global and epistemic metacognitive lens. To begin to fill these gaps, I present in Chapter 9 a case study independent of my phenomenographic study, examining one student's experience participating in a biology education research project. By considering her research papers, written reflections, and self-analysis, I follow the changes in her EMK about knowledge, how these changes affected her procedural skills, her scientific agency, and factors that influenced these changes.

8.4. Implications for Research

There are multiple implications for researchers interested in studying student epistemology in the biology discipline. While prior work primarily investigated student biology epistemologies through quantitative survey analyses, this research contributes an outcome space that qualitatively describes the epistemic practices students employed in a CURE through analysis of their arguments. The outcome space serves as a starting point for researchers to investigate epistemic practices that are productive in the discipline of biology, and to determine specific educational contexts that encourage students to apply these productive epistemic practices.

This study also advances the use of the Epistemic Thinking framework (Barzilai & Zohar, 2014, 2016) in the study of student epistemologies in biology. Participants' choice of epistemic practices was influenced by their EMK about knowledge in the context of the CURE as well as their EMS, specifically their reflexive processes. Further research should be conducted to investigate the interactions between epistemic thinking, reflexivity, and how this interaction affects the development of student biology epistemologies.

Finally, the methods used in this study advance the use of phenomenological methods in biology education research. The methods in this dissertation builds on the "blenderizing" technique as described in Pfirman (2018) to include "warp and weft" stages where I analyzed participant responses horizontally across interview questions (warp), then vertically through each individual participant (weft). This technique allowed me to gain a deeper perspective on the analysis by first getting an overview of the diversity of epistemic practices participants discussed for each interview question, followed by an analysis of how each participant applied these epistemic practices as they navigated the research process. Warp and weft was aided by constructing analysis memos as described in Lee *et al.* (2019), which served to contextualize interview excerpts within each participant's experience. These analysis memos were precursors to the three cases presented in Chapter 7, which contributed to further insights about the reasons students chose to apply specific epistemic practices in particular contexts. It is my hope that researchers find that the methods presented in this study serve as a useful guide for implementing phenomenography research.

9. CONCLUSIONS AND FUTURE WORK

9.1. Conclusions

In this work, I conducted a qualitative study that employed phenomenography to construct categories of description and an outcome space to understand the different ways in which students approached knowledge construction in a CURE. I chose to specifically study the epistemic practices students used to build arguments because argumentation is an epistemic and a core scientific process. While discussing their arguments, students described the epistemic practices of gathering information, checking credibility of information, constructing claims, and supporting experimental claims. Participants described these practices along two continuums: from a local (classroom) to global (outside-classroom) perspective, and from a focus on epistemic cognition (EC; the validity of information) to a focus on epistemic metacognition (EM; how the information was constructed). Students' epistemic practices placed along these two continuums, visualized as orthogonal axes, comprise the outcome space. Participants used these practices as they built four separate arguments in their CURE: 1) the instructor-guided hypothesis is reasonable, 2) the student-generated hypothesis is reasonable, 3) I accept or reject my hypothesis, and 4) the conclusions I have drawn are appropriate.

The outcome space revealed that participants applied epistemic practices differently depending on context, supporting and extending previous research describing epistemology as context-dependent. Participants discussed practices primarily in the EC/local quadrant when describing the instructor-guided hypothesis. However, discussions about these practices became comparatively more prevalent in the global and EM quadrants when participants discussed the student-generated hypothesis. As participants discussed epistemic metacognition, their focus shifted to include processes such as constructing knowledge. It is therefore important to encourage students to discuss their practices with epistemic metacognition, since the contemporary goals of biology education include teaching biology through the scientific practice.

During their interviews, students also discussed dissonance between their classroom perspectives and their perspectives on how professional science constructs knowledge. These reflections led to student descriptions of their beliefs about biological knowledge, and are examples of student engagement with reflexivity, the internal dialog wherein students evaluate and re-evaluate their actions and claims. This internal dialog leads to desirable epistemic ends such as the construction of valid knowledge and deep understanding. These insights illustrate how epistemic classroom discussions can contribute to student learning.

9.2. Future work

While the outcome space presented in this dissertation describes the epistemic practices that participants use to construct arguments, future work should include investigations into how students acquire these epistemic practices, why they choose to apply specific practices in particular contexts, and how instructors can support student use of productive epistemic practices.

During the analysis of my data, I also saw some connections between epistemology and concepts that were outside the scope of my study, including aspects of motivation such as future time perspective (FTP) and perceived instrumentality (PI). A second area of future research would be to explore the connections between FTP, PI, and epistemology. Future time perspective refers to how perceptions of the future affect motivational goal setting in the present (Husman & Lens, 1999), and perceived instrumentality refers to an individual's perceived usefulness of a task based on their future self (Kauffman & Husman, 2004). An example of FTP and PI interacting with epistemic practices in this study would be Joe, who applies epistemic practices that align with the biology discipline because he sees his future self as a botany researcher. This FTP motivates Joe to choose specific epistemic practices so he can build the skills he perceives as important for his future career.

A third area of proposed research is to study how students reconcile their strongly held antecedent epistemologies with the science epistemology presented in the classroom. Hannah provides an excellent example of how her antecedent epistemology clashed with the science epistemology of the CURE curriculum in this study. At the end of all of my interviews, I asked my participants if they had anything else to say that might add to my project. Hannah felt it was important to highlight her strong religious beliefs.

> I think that it might be important to note that there are probably beliefs that I have outside the scope of this course that affect like how I think about the reliability of scientific data and like theories I guess, so foundational science. Yeah I don't know I guess that's probably, that might be important.

> So I talked about evolution, so I'm a Christian and therefore like those beliefs that I have might cloud my judgment as to whether I consider the theory of evolution knowledge. To someone else who doesn't have that belief then maybe the Theory of Evolution is knowledge. So I don't know I guess that is important.

Hannah’s discussion of her strong religious beliefs seemed to affect her thoughts about what counts as knowledge. During her interview, Hannah stated that she considered measurements knowledge because they were objective, but she did not consider theories knowledge because they were interpretations. One possible explanation for her statements may be that she is attempting to reconcile her religious beliefs with her science epistemology. If she casts scientific theories as uncertain, she can retain her religious beliefs and at the same time accept scientific theories that may be in conflict with these beliefs. Studying the interactions between religious beliefs and epistemology may yield insights into public acceptance of science. Hannah’s case also brings in questions about how students’ antecedent cultures with the culture of the classroom. Do students integrate these cultures, or keep them separate? How do they reconcile cultures that may seem at odds, and how does this affect students’ science identities, especially identities of students that are part of traditionally underrepresented minority groups in STEM? I hope to continue investigating these questions in my future work.

APPENDICES

Appendix A: Required IRB Documents

Consent Form

Information about Being in a Research Study
University of Wisconsin-Madison and Clemson University

Student perceptions of what counts as knowledge in a course-based undergraduate research experience

Principle Investigator:
Lisa Benson, Ph.D.
Department of Engineering and Science Education
Clemson University
864-656-0417
lbenson@clemson.edu

Description of the Study and Your Part in It
You are invited to participate in a research study assessing Biocore students' perceptions of what counts as knowledge in biology. We are studying this through surveys, student work and interviews.

Your participation will involve providing access to an ungraded version of your written assignments (which may include posters, notes, or research proposals), and completing one survey online through Qualtrics. The survey takes about 10 minutes to complete. You may also be invited to participate in at least one 1-hour audio recorded interview about your research proposal.

Risks and Discomforts
If you are participating in the interview, we will ask you questions regarding your comfort, confidence level, personal opinions, or beliefs about your written assignments. Please do not share any information that may be sensitive or make you uncomfortable. You may refuse to answer or leave the interview at any time if you become uncomfortable.

Possible Benefits
There are no predicted personal benefits to your participation in this study. However, your information may benefit future students by helping us make research programs more effective for undergraduate students at UW-Madison and Clemson University. In addition, this research will be disseminated so that students and faculty at other institutions may benefit as well.

Incentives
If you are invited to participate in the interview(s), you will be given a $20 gift card after completion of each interview.

Protection of Confidentiality
We will do everything we can to protect your confidentiality. Your name will not be recorded in any way in the compiled interview or course work data. Your responses will be marked with a code. Only approved personnel will have the key which links your identity to that code, and this key will be destroyed as soon as all data have been collected and compiled. Additionally, current Biocore instructors are not involved in the recruitment of participants for this study. As such, they will not know the identity of participants in this study, nor will they have access to the data until after the course has been completed. Once the course is completed, Biocore instructors will only have access to deidentified data. Audio recordings and other electronic data will be stored on a password-protected database accessible only by the researchers apporved on this study. Audio data will be stored until it has been transcribed. Once transcribed, the audio data will be destroyed. Your identity will not be revealed in any presentation or publication that might result from this study.

Choosing to Be in the Study
You may choose not to participate in any part of the study and you may withdraw your consent to participate at any time. Additionally, your decision whether or not to participate in the research will not affect your grades or undergraduate career.

Exclusion Requirements
Participants must be at least eighteen years of age to be eligible to participate.

Contact Information
If you have any questions or concerns about this study or if any problems arise, please contact Lisa Benson at Clemson University at 864-656-0417.

If you have any questions or concerns about your rights in this research study, please contact the Clemson University Office of Research Compliance (ORC) at 864-656-0636 or irb@clemson.edu. If you are outside of the Upstate South Carolina area, please use the ORC's toll-free number, 866-297-3071.

Consent

I have read this form and have been allowed to ask any questions I might have. My signature below signifies that I agree to take part in this study.

Participant's signature: ____ Date: ________________

Printed Name: ______________________________

E-mail address: _____________________________

An electronic copy of this form will be provided to you.

Announcement Script to be presented to Students

Hi, My name is ________________

I am Ph.D. Candidate in the Engineering and Science Education Department at Clemson University, and we are conducting a study to investigate undergraduate Biocore students' views on knowledge and research. You are invited to participate in an online survey related to this topic if you are over 18 years of age and are enrolled in a Biocore class. During the semester, I will also ask you to provide me with electronic copies of your ungraded assignments, and invite students to participate in one or more individual 1-hour interviews. You will be compensated with a $20 gift card after each interview.

The results of this study will be used to understand how students view their research and classroom learning experiences and to inform future programs. We would greatly appreciate your participation, as your perspective on classes and research is valuable to this study. If you are interested in participating in this study, please sign the consent form that I have provided. If you have any questions, I would be happy to answer them. Thank you.

Initial Recruitment Survey

Need For Closure Questions for Participant Selection

Read each of the following statements and indicate how much each statement reflects your attitudes, beliefs, and experience. It is important that you know that there are no “right” or “wrong” answers to these questions.

Q2a = I don't like situations that are uncertain.

Q2b = I prefer interacting with people whose opinions are very different from my own.

Q2c = Even after I've made up my mind about something, I am always eager to consider a different opinion.

Q2d = I dislike questions which could be answered in many different ways.

Q2e = I feel uncomfortable when someone's meaning is unclear to me.

Q2f = It's annoying to listen to someone who cannot seem to make up his/her mind.

Q2g = When considering most conflict situations, I can usually see how both sides could be right.

Q2h = I do not usually consult many different opinions before forming my own view.

Q2i = I always see many possible solutions to problems I face.

Q2j = I feel uncomfortable when I don't understand the reason why an event occurred in my life.

Q2k = When I am confused about an important issue, I feel very upset.

Q2l = I'd rather know bad news than stay in a state of uncertainty.

Q2m = When thinking about a problem, I consider as many different options on the issue as possible.

Q2n = I like to know what people are thinking all the time.

Q2o = In most social conflicts, I can easily see which side is right and which is wrong.

VALUES: 0 through 6 (rating scale); 0 = “Not at all”, 6 = “Very much so”

General Recruitment Questions

1. I am 18 years of age or older: yes/no
2. Are you willing to participate in a 1-hour interview? Yes/no
3. (if yes to interview) Please provide an e-mail address so we can contact you about an interview appointment. (open)

E-mail Recruitment Script for Interview

Subject: You are invited to participate in an interview to study the effects of Biocore instruction on student reasoning

Dear [Insert Name Here]

We would like to invite you to participate in an interview. If you are still willing to complete a 1-hour interview on your experience in your Biocore class, please sign up for an interview appointment here. At the completion of the interview, you will receive a $20 incentive (Amazon card) as a thank you for your contribution to our work.

During the interview, you will be asked several open-ended questions relating to your experience in your Biocore class. Your responses will be audio recorded and transcribed to written text for analysis. Your responses will help us understand how your experiences may be influencing your views on knowledge and where it comes from. We will protect your privacy and personal information by omitting identifying information (yourself or others by name) from our analysis.

Please schedule a time to meet by clicking on the following link:

https://biocore_interviews.youcanbook.me/

This interview is part of a research study conducted by Dr. Lisa Benson. If you have any questions or concerns about this study or if any problems arise, please contact Dr. Lisa Benson at Clemson University at 864.656.0417 or lbenson@clemson.edu. If you have any questions or concerns about your rights in this research study, please contact the Clemson University Office of Research Compliance (ORC) at 864-656-0636 or irb@clemson.edu. If you are outside of the upstate South Carolina area, please use the ORC's toll-free number, 866-297-3071.

Thank you in advance for your participation!

Interview Protocol

Pre-interview

- Explain consent form, ask student if he/she is over 18, have student sign
- Ask student if it is OK for this interview to be audio recorded
- Start audio recording on two recording devices

Interview Questions (Possible follow-up questions indented)

1. What is the importance of this project to you?
    a. What importance does this study have to you?
    b. What is the take-home message you want to get across in your poster?
2. How did you come up with your hypothesis? Where did you get information for your hypothesis?
    a. How did you get this information?
    b. Why do you trust this information?
    c. Why do you trust this information source?
3. What conclusions did you draw? How did you know they were correct?
    a. How does this information connect to the conclusion(s) that you make?
    b. How is this conclusion supported?
4. Are there parts of your project that you would consider knowledge? What makes those components knowledge?

Demographic questions

1. Are you a first-generation college student?

2. What ethnicity and/or race do you identify with?
3. What gender do you identify with?
4. What is your college major?
5. What year are you?

Appendix B: Research Team Positionality Statements

Cazembe

How do I know something is true in research?
This comes from much fact checking and then determining what can be considered "true." One important thing to science and the research involved in science is that many "conclusions" are not meant to definitively say something is true for all instances or situations. It may have been proven under certain conditions and verified by multiple people, which makes us/me more comfortable and confident in believing it, but that does not make it an absolute fact. With that being said, many times when writing or referring to other research done, I attempt to caution and state that this is what prior work has shown, and it helps to strengthen my case in these ways, however, I am more so saying that I/we are trusting the research and work that other researchers have done because it seems reasonable based on what is known about how research or GOOD research is conducted. So cross checking research to see if there is evidence to support or refute it, examining the validity of that evidence as well as the research method based on what I know about how good research should be conducted, and being careful to frame the "truth" or validity of that research within the confines of the point/points I'm trying to make are the ways that I know/determine that something in research is true.

What are "bad" epistemologies in your opinion?
I tend to believe the extremes of thought/opinions are when you see "bad" cases. So people who would think very absolutist and that everything must be binary, black or white, have a definitive answer, etc. would have a bad epistemic thought process in my mind. Similarly, people who thought that there is nothing that can ever be considered knowledge and it is all subjective and up for interpretation also have bad mindsets. The latter is how you get into the "BS Zone" (Trademark pending). Both very objective or very subjective can lead to logical fallacies that stop...more ideal knowledge building. Also, I think people unable to construct their own knowledge run the risk of having bad epistemologies. It can lead to being easily influenced by any type of authority figure or trusted source of knowledge and I do not think it is a good way to go about thinking. Again, the foil to that of thinking your opinion or way of thinking is automatically more valuable or valid than other ones, even when they are authoritative or trusted sources on the knowledge being discussed would be a bad epistemic practice.

Jason

How do I know something is true in research?

It is difficult to know what is TRUE, since reality is different from each observer's lens. As a qualitative phenomenographic researcher, I am tasked with identifying phenomena and, for this project, modeling students' epistemic cognition. Something is true for us, in my opinion, as a research team when we agree on what students are saying in the data, it is consistent with other passages coded with the same category, and when this research has been done with an open-mind and with as little bias as possible. So I would say I can't KNOW that something is true, but I follow the guidelines of the methodology and do everything I can to represent the data as well as possible. I can be true to the method and true to our categorical codes and my own ethical standards of research.

What are "bad" epistemologies?

I have been told enough times that there are no "bad" epistemologies to the point where I think I mostly believe that. And yet, as a person who values open-mindedness, honesty and transparency as paramount, I can't help but look at a person poorly who expresses close-minded thoughts, ideas, or opinions. Without the ability to challenge a participant, since I am only looking at their recorded words, I am likely to become frustrated with a student who is particularly dismissive or aloof or sounds like they are trying to placate the researcher rather than answer the questions seriously and honestly. I have found that if a student is transparent about them not knowing something, I look kindly upon that. But if a student seems, in my view of course, to know what is going on but fumbles around an answer, I naturally question the validity of what they are saying and less likely to give them credit for the epistemic cognitions stated (is it truly epistemic if it's not real? Does it matter? I don't know the answers to these questions).

Appendix C: Example of Analytical Memo

Cat Q1

The importance of this project for Cat really is learning skills. She learned how to make a poster, how to think deeply about the project, and how to explain science.

Just learning to like really think more in depth and like look at every aspect of a project from every possible perspective because you never know which perspective is the most effective of explaining the question that you're asking. There's also a lot of like looping around when thinking about a question. It's like you just like go through the same cycles but just with slight variation.

Cat Q2

The hypothesis for tier 1 was standard, while for tier, the hypothesis was that growing many plants in a container would cause stress and increase API. Cat is quite straightforward about doing things because she wants a good grade. She says point blank that she trusts the lab manual because her grade depends on it. On the other hand, she trusts journal articles because they're logical, are written by professionals and come from universities.

> Because they just had very logical, well explained methods that produced results and I trust that people writing papers usually have pretty good understanding of what implications can be drawn from those results. I guess it's just trust and experience […] A lot of the journals were from like specific universities. I can't remember which but it was pretty apparent that it was research conducted through universities.

The University thing comes up with Cat too. When I asked her to go through her BR, she started going into her results, saying she didn't really feel like the methods addressed the question.

> What we didn't really factor in is the plants in their containers were saturated with a nutrient solution so there wasn't really partitioning of resources because the entire thing was constantly saturated. The stress induced by water loss or having to share nutrients, it wasn't there exactly. It was literally just root density and because the Brassica Rapa Fast plants were bred to not mind partitioning of light and close proximity, it didn't have as much of an impact as we might have hoped. […] Since Fast plants were bred not to care about that, then it didn't really induce any stress.

We'll talk more about the results later. In the meantime, we had a discussion about the importance of the BR. Cat feels that the BR helps to improve the credibility of the results from the study. She explains below.

> It gives a reason for why the thought would happen and it's important because it shows that you do really know this topic, you've been researching what you're trying ... you've been researching background information on the question that you're approaching. It gives you a little bit more credibility to have a really good biological rationale and you can trust your results more if you have a good biological rationale. […] It makes results a little bit more trustworthy because it shows that other people have found potential explanations for your question so it almost eliminates the possibility that results are a fluke or anything like that.

She points at a mechanistic explanation that the BR provides for their avenue of thinking. Additionally, she feels that the BR shows others that you’ve done research and know the topic well. The results are also more trustworthy because you can compare your results to those that have already been published.

When she explains how they got to the tier 1 hypothesis, I see a connection to her feeling about getting a good grade in class. This roundabout conversation she’s talking about. It’s like she knows what’s up, and is fully in the classroom mode here, going through the motions. They’re trying to make us do this, think this way. I’ll play along. Like what Berland says in the epistemology in practice paper.

> We talked about it a lot in discussion, like it was one of those roundabout conversations where we were really looking for specific keywords in talking about this. It makes sense logically because the actual allelic variations are reduced to an extent to only have the like more purple alleles within the population but since it is a quantitative trait you can't know that for sure, that all of the allelic variation for the green or like lighter colored API aren't still in the population of the new seeds if that makes sense.
>
> Yeah, so we talked a lot about like the graph of what the parental generation looked like in terms of it's API and how if you like slice off the top 20% and how his has a more condensed group of allelic variation, but you still have a lot of different alleles that will code for a lighter color of API and we just talked about oh, how does that compare to a 1% sliver or even a 5% sliver type thing. And just talking about the different types of crosses that we could arise at because it's a quantitative trait and there is a lot of different variables obviously.

She seems to be playing it cool about learning all this stuff but she really gets it. Whether she likes it or not, she’s looking at mechanistic explanation of the hypothesis. Almost like an involuntary epistemology? Lol. Tier 2 hypothesis formation is less complex. It’s an assumption based on similar observations. Other forms of stress increase API, so this overcrowding, which should cause stress, should also increase API.

> We just saw that other different types of stressors acting on Brassica Rapa or even monoculture plants would just induce stress and so we said oh, maybe root

> density is enough of a stressor to also induce a higher API as shown in other studies of Brassica plants.

Cat Q3

Cat rejected both hypotheses, something she was surprised about. It seems that the "I'm doing it for the grade" was still hiding some knowledge from authority. You can see this break in her armor in the excerpt below.

> It was very interesting to like almost discard the faith in what the teacher said just based on my results and it's like oh, were my results really accurate with such a small sample size or should I have gotten a difference in the phenotype? […] I think I had to trust the science because other groups also got those results so adding to the sample size increased confidence. I'd say yeah, it's nice to like see it with my own eyes what the teachers are saying versus what they're not saying.

So here she's starting to form a perception that the instructors are not always correct, or maybe that the instructors do not always give you the "right answer." She also supports this conclusion with some consensus with the rest of the class. Other students got these results too, which adds to her confidence. However, she is careful to temper her confidence.

> I just know that one other group I compared with, they had the same results as we did so that doubled my own group sample size... So it doubled my confidence.

It looks like she sees nuance in the conclusions she makes. We see more of this when we talk about her tier 2 results. First, she explains that their assumption probably was not correct because of what a professor told them.

> One of the professors said that he tried growing a whole bunch of Brassica Rapa in like a little film bottle and he could fit about 30 of them in there and they weren't doing well but they survived, so it's possible that we didn't actually cram them in enough. We thought that cramming them in four times the amount of the control should have had some impact at least but we didn't find that.

Then comes the interesting part:

> Overlap of the error bars means that the difference isn't so drastic that you can draw a conclusion from it saying that it is different or it means that if you do try and draw a conclusion, you're not going to be very confident in it, in saying oh, this is definitely because of this factor I'm proposing but instead it could have happened randomly.

She doesn't use the overlapping error bars as an absolute. You can draw conclusions, but you're just not confident about it. Nuance.

Cat Q4

Learning, just as in question 1, is in her definition of knowledge. When I asked if she considered anything in her project knowledge, this is how she answered:

> I would. I learned a lot about needing to know the actual biological mechanisms that occur behind these differences and I learned a lot more about what it means to really understand what's going on in a system because in my biological rationale and background, I didn't include anything about genes or things ... like very specific things, they're all a lot more general but when I was working on the group project we had like tons of specific things and like talking about oh, this gene produces this and this one this and like everything like that.

The specificity she's talking about is mechanistic reasoning. This one connects to that one and so on. She's explaining how she got to her biological rationale, the proposed mechanism behind her hypothesis. She also talks about skill attainment. This is an important aspect of knowledge for her. It's something she didn't know before.

> Honestly, the biggest thing I took away from this project is like knowledge on how to actually perform a project like this and a lot more knowledge on the actual experimental process, so I think the results of this poster isn't as important as the fact that I know how to do it now.

Cat also talks about understanding as knowledge. To her, understanding is looking at something from many different angles.

> What it means to understand something is be able to look at a system or a question from multiple perspectives and to finally reach a conclusion that's like a culmination of all these different ideas in a way that's like logical and actually representative of what's going on I suppose. With the Brassica Rapa experiment in particular, you can look at it from the biological mechanism where it's talking about like genes producing, but you can also look at it from environmental impacts where it's like the stressors are going to increase anthocyanin production. You can also look at it from like the allelic variation so you have to look at all the different perspectives to really understand something and gain knowledge about it.

Looking at it from all these perspectives helps you gain knowledge. Finally, Cat differentiates different kinds of knowledge: stuff she didn't know before, and scientific knowledge. She doesn't feel like she constructed scientific knowledge in this project.

> I would say just because I don't think the conditions of the experiment were extreme enough, I don't think much scientific knowledge would have been gained by this project. I think if you could definitely say that root density didn't have an

> impact, then I think you would gain scientific knowledge, but I don't think that these methods can prove that.

It seems like the fact that she can't make strong conclusions from her data makes her feel that she did not make scientific knowledge.

Appendix D: Poster Analysis Template

Template Poster analysis

Hypothesis 1:

Data:

Warrant:

Hypothesis 2:

Data:

Warrant:

Interview Questions

Hypothesis 1:

- How is this hypothesis supported?
- How does your biological rationale contribute to your hypothesis?

Hypothesis 2:

- How is this hypothesis supported?
- How does your biological rationale contribute to your hypothesis?

Expected/Alternative results

- Why did you include this section in your poster?
- How did you come up with these results? How do you know that they are correct?

Methods:

REFERENCES

Akerlind, G. (2005). Learning about phenomenography: Interviewing, data analysis and the qualitative research paradigm. In J. Bowden & P. Green (Eds.), *Doing Develpmental Phenomenography* (pp. 63–73). Melbourne, Australia: RMIT University Press.

Akerlind, G., Bowden, J., & Green, P. (2005). Learning to do Phenomenography: A Reflective Discussion. In J. Bowden & P. Green (Eds.), *Doing Developmental Phenomenography* (pp. 74–103). Melbourne, Australia: RMIT University Press.

Åkerlind, G. S. (2012). Variation and commonality in phenomenographic research methods. *Higher Education Research & Development*, *31*(1), 115–127. https://doi.org/10.1080/07294360.2011.642845

Archer, M. S. (2012). *The Reflective Imperative in Late Modernity*. Cambridge, UK: Cambridge University Press.

Associates, H. R. (2018). Fulfilling the American dream: Liberal education and the future of work: Selected findings from online surveys of business executives and hiring managers. *Association of American Colleges and Universities*, (July), 20. Retrieved from https://www.aacu.org/sites/default/files/files/LEAP/2018EmployerResearchReport.pdf

Auchincloss, L. C., Laursen, S. L., Branchaw, J. L., Eagan, K., Graham, M., Hanauer, D. I., … Dolan, E. L. (2014). Assessment of course-based undergraduate research experiences: Ameeting report. *CBE Life Sciences Education*, *13*(1), 29–40. https://doi.org/10.1187/cbe.14-01-0004

Bangera, G., & Brownell, S. E. (2014). Course-based undergraduate research experiences can make scientific research more inclusive. *CBE Life Sciences Education*, *13*(4). https://doi.org/10.1187/cbe.14-06-0099

Barzilai, S., & Chinn, C. A. (2017). On the Goals of Epistemic Education: Promoting Apt Epistemic Performance. *Journal of the Learning Sciences*, *27*(3), 1–37. https://doi.org/10.1080/10508406.2017.1392968

Barzilai, S., & Zohar, A. (2014). Reconsidering personal epistemology as metacognition: A multifaceted approach to the analysis of epistemic thinking. *Educational Psychologist*, *49*(1). https://doi.org/10.1080/00461520.2013.863265

Barzilai, S., & Zohar, A. (2016). Epistemic (Meta)Cognition: Ways of Thinking About Knowledge and Knowing. In J. A. Greene, W. A. Sandoval, & I. Braten (Eds.), *Handbook of Epistemic Cognition* (pp. 409–424). New York, NY.

Baxter Magolda, M. B. (1992). *Knowing and reasoning in college: Gender-related patterns in students' intellectual development.* San Francisco, CA: Jossey Bass.

Belenky, M. F., Clinchy, B. M., Goldberger, N. R., & Tarule, J. M. (1983). *Women's ways of knowing: The development of self voice and mind*. New York, NY: Basic Books.

Bell, R. L., & Lederman, N. G. (2003). Understandings of the Nature of Science and

Decision Making on Science and Technology Based Issues. *Science Education*, *87*(3), 352–377. https://doi.org/10.1002/sce.10063

Bell, R. L., Matkins, J. J., & Gansneder, B. M. (2011). Impacts of contextual and explicit instruction on preservice elementary teachers' understandings of the nature of science. *Journal of Research in Science Teaching*, *48*(4), 414–436. https://doi.org/10.1002/tea.20402

Berland, L. K., Schwarz, C. V., Krist, C., Kenyon, L., Lo, A. S., & Reiser, B. J. (2016). Epistemologies in practice: Making scientific practices meaningful for students. *Journal of Research in Science Teaching*, *53*(7). https://doi.org/10.1002/tea.21257

Bowden, J. (2005). Reflections on the phenomenographic team research process. In J. Bowden & P. Green (Eds.), *Doing Develpmental Phenomenography* (pp. 11–31). Melbourne, Australia: RMIT University Press.

Bowden, J. a. (2000). The nature of phenomenographic research. *Phenomenography*, 1–18.

Bowden, J. a, & Green, P. (2005). *Doing developmental phenomenography*. *Qualitative research methods*. Retrieved from http://search.informit.com.au/browsePublication;isbn=1921166037;res=E-LIBRARY

Brewer, C., & Smith, D. (2011). *Vision and change in undergraduate biology education: a call to action*. Washington, DC. https://doi.org/10.1002/(SICI)1098-2736(200004)37:4<295::AID-TEA2>3.0.CO;2-2

Brownell, S.E., & Kloser, M. J. (2015). Studies in Higher Education Toward a conceptual framework for measuring the effectiveness of course-based undergraduate research experiences in undergraduate biology. *Studies in Higher Education*, *40*(3), 525–544. https://doi.org/10.1080/03075079.2015.1004234

Brownell, Sara E, Hekmat-scafe, D. S., Singla, V., Seawell, P. C., Imam, J. F. C., Eddy, S. L., … Cyert, M. S. (2015). A High-Enrollment Course-Based Undergraduate Research Experience Improves Student Conceptions of Scientific Thinking and Ability to Interpret Data. *CBE-Life Sciences Education*, *14*, 1–14. https://doi.org/10.1187/cbe.14-05-0092

Charmaz, K. (2014). *Constructing Grounded Theory* (2nd ed.). Los Angeles, CA: SAGE Publications.

Chinn, C. A., & Rinehart, R. (2016). Epistemic Cognition and Philosophy: Developing an New Framewrok for Epistemic Cognition. In J. A. Greene, W. A. Sandoval, & I. Braten (Eds.), *Handbook of Epistemic Cognition* (pp. 460–478). New York, NY: Routledge.

Chinn, C., Rinehart, R., & Buckland, A. (2014). Epistemic cognition and evaluating information: Applying the AIR model of epistemic cognition. In D. Rapp & J. Braasch (Eds.), *Processing inaccurate information: Theoretical and applied perspectives from cognitive science and the educational sciences* (pp. 425–453). Cambridge, MA: MIT Press.

Chinn, Clark a., Buckland, L. a., & Samarapungavan, A. (2011). Expanding the Dimensions of Epistemic Cognition: Arguments From Philosophy and Psychology. *Educational Psychologist*, *46*(3), 141–167.

https://doi.org/10.1080/00461520.2011.587722
Creswell, J. (2012). *Qualitative Research & Design: Choosing Among Five Approaches* (Third Edit). Thousand Oaks, CA: SAGE Publications.
Driver, R., Learch, J., Millar, R., & Scott, P. (1996). *Young People ' s Images of Science*. Buckingham, UK: Open University Press. https://doi.org/10.1088/0031-9120/31/3/023
Duschl, R. (2008). Science Education in Three-Part Harmony: Balancing Conceptual, Epistemic, and Social Learning Goals. *Review of Research in Education*, *32*(1), 268–291. https://doi.org/10.3102/0091732X07309371
Duschl, R. A., & Osborne, J. (2002). Supporting and Promoting Argumentation Discourse in Science Education. *Studies in Science Education*, *38*(1), 39–72. https://doi.org/10.1080/03057260208560187
Elby, Andrew, & Hammer, D. (2001). On the substance of a sophisticated epistemology. *Science Education*, *85*(5), 554–567.
Elby, Andrew, Macrander, C., & Hammer, D. (2016). Epistemic Cognition in Science. In J. A. Greene, W. A. Sandoval, & I. Braten (Eds.), *Handbook of Epistemic Cognition* (pp. 113–127). New York, NY: Routledge.
Elby, Andy, & Hammer, D. (2010). *Epistemological resources and framing: a cognitive framework for helping teachers interpret and respond to their students' epistemologies*. (L. D. Bendixen & F. C. Feucht, Eds.), *Personal Epistemology in the Classroom: Theory, Research, and Implications for Practice*. New York, NY: Cambridge University Press. https://doi.org/10.1017/CBO9780511691904.013
Faber, C., Kajfez, R., Lee, D., McAlister, A., Ehlert, K., Kennedy, M., & Benson, L. (2020). Undergraduate Engineering Students' Perceptions of Research and Researchers. *Journal of Engineering Education*.
Ferrario, C. G. (2004). Developing Clinical reasoning Strategies: Congnitive Shortcuts. *Journal for Nurses in Staff Development*, *20*(5), 229–235.
Flavell, J. H. (1979). Metacognition and Cognitive Monitoring: A New Area of Cognitive-Development Inquiry. *American Psychologist*, *34*(10), 906–911. https://doi.org/10.1037/0003-066X.34.10.906
Gottesman, A. J., & Hoskins, S. G. (2013). CREATE Cornerstone: Introduction to scientific thinking, a new course for STEM-interested freshmen, demystifies scientific thinking through analysis of scientific literature. *CBE Life Sciences Education*, *12*(1), 59–72. https://doi.org/10.1187/cbe.12-11-0201
Green, P. (2005). A rigorous journey into phenomenography: From a naturalistic inquirer standpoint. In J. a Bowden & P. Green (Eds.), *Doing developmental phenomenography* (pp. 32–46). Melbourne: RMIT University Press.
Green, P., & Bowden, J. a. (2005). Principles of Developmental Phenomenography. *Malaysian Journal of Qualitative Research*, *2*, 52–70.
Guetzkow, H. (1950). Unitizing and categorizing problems in coding qualitative data. *Journal of Clinical Psychology*, *6*(1), 47–58.
Hallman, W. K. (2017). What the public thinks and knows about science-and why it matters. *The Oxford Handbook of the Science of Science Communication*, (June 2019), 61–72. https://doi.org/10.1093/oxfordhb/9780190497620.013.6

Hatano, G., & Inagaki, K. (1986). Two Courses of Expertise. In H. Stevenson, H. Azuma, & K. Hakuta (Eds.), *Child Development and Education in Japan* (pp. 27–36). New York, NY: W. H. Freeman.
Hofer, B K. (2004). Epistemological Understanding as a Metacognitive Process: Thinking Aloud During Online Searching. *Educational Psychologist*, *39*(1), 43–55. https://doi.org/10.1207/s15326985ep3901_5
Hofer, Barbara K, & Pintrich, P. R. (1997). The Development of Epistemological Theories : Beliefs about Knowledge and Knowing and Their Relation to Learning. *Review of Educational Research*, *67*(1), 88–140.
Hoskins, S. G., Lopatto, D., & Stevens, L. M. (2011). The C.R.E.A.T.E. approach to primary literature shifts undergraduates' self-assessed ability to read and analyze journal articles, attitudes about science, and epistemological beliefs. *CBE Life Sciences Education*, *10*(4), 368–378. https://doi.org/10.1187/cbe.11-03-0027
Hunter, A., Laursen, S. L., & Seymour, E. (2007). Becoming a scientist: The role of undergraduate research in students' cognitive, personal, and professional development. *Science Education*, *91*(1), 36–74. https://doi.org/10.1002/sce
Husman, J., & Lens, W. (1999). The Role of the Future in Student Motivation. *Educational Psychologist*, *34*(2), 113–125. https://doi.org/10.1207/s15326985ep3402_4
Jiménez-Aleixandre, P. M., Bugallo Rodrìguez, A., & Duschl, R. A. (2000). " Doing the Lesson " or " Doing Science " : Argument in High School Genetics ARGUMENT. *Science Education*, *84*, 757–792.
Kauffman, D. F., & Husman, J. (2004). Effects of time perspective on student motivation: Introduction to a special issue. *Educational Psychology Review*, *16*(1), 1–7. https://doi.org/10.1023/B:EDPR.0000012342.37854.58
Kelly, G. J. (2008). Inquiry, activity and epistemic practice. In RA Duschl & R. Grandy (Eds.), *Teaching Scientific Inquiry: Recommendations for research and implementation* (pp. 99–117). Rotterdam: Sense Publishers.
Kenyon, K. L., Onorato, M. E., Gottesman, A. J., Hoque, J., & Hoskins, S. G. (2016). Testing CREATE at community colleges: An examination of faculty perspectives and diverse student gains. *CBE Life Sciences Education*, *15*(1), 1–19. https://doi.org/10.1187/cbe.15-07-0146
King, P. M., & Kitchener, K. S. (1994). *Developing reflective judgment: Understanding and promoting intellectual growth and critical thinking in adolescents and adults*. San Francisco, CA: Jossey Bass.
Kinnunen, P., & Simon, B. (2012). Phenomenography and grounded theory as research methods in computing education research field. *Computer Science Education*, *22*(2), 199–218. https://doi.org/10.1080/08993408.2012.692928
Knorr-Cetina, K. (1999). *Epistemic Cultures: How the Sciences Make Knowledge*. Cambridge, MA: Harvard University Press.
Kruglanski, A. W. (1990). Lay Epistemic Theory in Social-Cognitive Psychology, *1*(3), 181–197.
Kuhn, D. (1991). *The Skills of Argument*. New York, NY: Cambridge University Press.
Kuhn, D. (1992). Thinking as argument. *Harvard Educational Review*, *62*(2), 155–178.

https://doi.org/10.17763/haer.62.2.9r424r0113t670l1
Kuhn, D. (1999). A developmental model of critical thinking. *Educational Researcher*, *28*(2), 16–46. https://doi.org/10.3102/0013189X028002016
Lave, J., & Wenger, E. (1991). *Situated Learning: Legitimate peripheral participation*. Cambridge, UK: Cambridge University Press.
Lederman, Norm G., Abd-El-Khalick, F., Bell, R. L., & Schwartz, R. S. (2002). Views of Nature of Science Questionnaire: Toward Valid and Meaningful Assessment of Learners' Conceptions of Nature of Science. *Journal of Research in Science Teaching*, *39*(6), 497–521. https://doi.org/10.1002/tea.10034
Lederman, Norman G. (1992). Students' and teachers' conceptions of the nature of science: A review of the research. *Journal of Research in Science Teaching*, *29*(4), 331–359. https://doi.org/10.1002/tea.3660290404
Lederman, Norman G. (2007). Nature of Science : Past , Present , and Future. In S. K. Abell & N. G. Lederman (Eds.), *Handbook of research on science education* (1st ed., pp. 831–880). New York, NY: Routledge.
Lee, D. M., McAlister, A. M., Ehlert, K., Faber, C. J., Kajfez, R. L., Creamer, E., & Kennedy, M. S. (2019). The Use of Analytical Memo Writing in a Mixed Methods Grounded Theory Study. In *Frontiers in Education (FIE) Conference Proceedings*. Cincinnati, OH.
Linn, M. C., Palmer, E., Baranger, A., Gerard, E., & Stone, E. (2015). Undergraduate research experiences: Impacts and opportunities. *Science*, *347*(6222), 1261757–1261757. https://doi.org/10.1126/science.1261757
Lombardi, D., Nussbaum, E. M., & Sinatra, G. M. (2016). Plausibility Judgments in Conceptual Change and Epistemic Cognition. *Educational Psychologist*, *51*(1). https://doi.org/10.1080/00461520.2015.1113134
Longino, H. E. (2000). Toward an epistemology for biological pluralism. In R. Creath & J. Maienschein (Eds.), *Biology and Epistemology* (pp. 261–286). New York, NY: Cambridge University Press.
Longino, H. E. (2002). *The Fate of Knowledge*. Princeton, NJ: Princeton University Press.
Lopatto, D. (2004). Survey of Undergraduate Research Experiences (SURE): first findings. *Cell Biology Education*, *3*(4), 270–277. https://doi.org/10.1187/cbe.04-07-0045
Lopatto, D. (2007). Undergraduate research experiences support science. *CBE-Life Sciences Education*, *6*, 297–306. https://doi.org/10.1187/cbe.07
Louca, L., Elby, A., Hammer, D., & Kagey, T. (2004). Epistemological Resources: Applying a New Epistemological Framework to Science Instruction. *Educational Psychologist*, *39*(1), 57–68. https://doi.org/10.1207/s15326985ep3901_6
Magnus, D. (2000). Down the primrose path: competing epistemologies in early twentieth-century biology. In R. Creath & J. Maienschein (Eds.), *Biology and Epistemology* (pp. 91–121). New York, NY: Cambridge University Press.
Marton, F. (1986). Phenomenography - a research approach to investigating defferent understandings of reality. *Journal of Thought*, *21*(3), 28–49. Retrieved from http://www.jstor.org/stable/42589189

Marton, F., & Booth, S. (1997). *Learning and Awareness*. Mahwah, NJ: Lawrence Earlbaum Associates, Inc.
Mylopoulos, M., Kulasegaram, K., & Woods, N. N. (2018). Developing the experts we need: Fostering adaptive expertise through education. *Journal of Evaluation in Clinical Practice*, *24*(3), 674–677. https://doi.org/10.1111/jep.12905
Mylopoulos, M., & Woods, N. N. (2017). When I say … adaptive expertise. *Medical Education*, *51*(7), 685–686. https://doi.org/10.1111/medu.13247
Nussbaum, E. M., & Sinatra, G. M. (2003). Argument and conceptual engagement. *Contemporary Educational Psychology*, *28*(3), 384–395. https://doi.org/10.1016/S0361-476X(02)00038-3
Nussbaum, E. M., Sinatra, G. M., & Poliquin, A. (2008). Role of epistemic beliefs and scientific argumentation in science learning. *International Journal of Science Education*, *30*(15), 1977–1999. https://doi.org/10.1080/09500690701545919
Olimpo, J. T., Fisher, G. R., & Dechenne-peters, S. E. (2016). Development and Evaluation of the Tigriopus Course-Based Undergraduate Research Experience : Impacts on Students ' Content Knowledge , Attitudes , and Motivation in a Majors Introductory Biology Course, 1–15. https://doi.org/10.1187/cbe.15-11-0228
Perry, W. G. (1970). *No Title* (1st ed.). New York, NY: Holt.
Pfirman, A. L. (2018). *Exploring Underrepresented Doctoral Students' Conceptualizations of the Student-Advisor Relationship in Chemistry*. *Dissertation*. Clemson University. https://doi.org/10.1002/(SICI)1097-4695(20000605)43
Qian, G., & Alvermann, D. (1995). Role of Epistemological Beliefs and Learned Helplessness in Secondary School Students' Learning Science Concepts From Text. *Journal of Educational Psychology*, *87*(2), 282–292. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-21844510692&partnerID=40&md5=38e4b22b3bd3b972935c5bc617fd1244
Rodenbusch, S. E., Hernandez, P. R., Simmons, S. L., Dolan, E. L., Services, H., Hughes, H., & Chase, C. (2016). Early Engagement in Course-Based Research Increases Graduation Rates and Completion of Science , Engineering , and Mathematics Degrees. *CBE-Life Sciences Education*, *15*, 1–10. https://doi.org/10.1187/cbe.16-03-0117
Rowland, S., Pedwell, R., Lawrie, G., Lovie-Toon, J., & Hung, Y. (2016). Do we need to design course-based undergraduate research experiences for authenticity? *CBE Life Sciences Education*, *15*(4). https://doi.org/10.1187/cbe.16-02-0102
Russ, R. S. (2014). Epistemology of Science vs . Epistemology for Science. https://doi.org/10.1002/sce.21106
Russell, C. B., & Weaver, G. C. (2011). A comparative study of traditional, inquiry-based, and research-based laboratory curricula: impacts on understanding of the nature of science. *Chem. Educ. Res. Pract.*, *12*(1), 57–67. https://doi.org/10.1039/C1RP90008K
Saldaña, J. (2016). *The Coding Manual for Qualitative Researchers* (3rd ed.). Los Angeles, CA: SAGE Publications.
Samarapungavan, A., Westby, E. L., & Bodner, G. M. (2006). Contextual epistemic development in science: A comparison of chemistry students and research chemists.

Science Education, *90*(3), 468–495. https://doi.org/10.1002/sce.20111
Sandoval, W. (2014). Science Education's Need for a Theory of Epistemological Development. *Science Education*, *98*(3), 383–387. https://doi.org/10.1002/sce.21107
Sandoval, W. A. (2005). Understanding students' practical epistemologies and their influence on learning through inquiry. *Science Education*, *89*(4), 634–656. https://doi.org/10.1002/sce.20065
Schommer-aikins, M., Duell, O. K., & Hutter, R. (2014). Beliefs , Mathematical Beliefs , and Academic Students, *105*(3), 289–304.
Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology*, *82*(3), 498–504. https://doi.org/10.1037/0022-0663.82.3.498
Semsar, K., Knight, J. K., Birol, G., & Smith, M. K. (2011). The colorado learning attitudes about science survey (class) for use in biology. *CBE Life Sciences Education*, *10*(3), 268–278. https://doi.org/10.1187/cbe.10-10-0133
Sinatra, G. (2016). Thoughts on Knowledge About Thinking About Knowledge. In J. A. Greene, W. A. Sandoval, & I. Braten (Eds.), *Handbook of Epistemic Cognition* (pp. 479–491). New York, NY: Routledge.
Sochacka, N. W., Walther, J., & Pawley, A. L. (2018). Ethical Validation: Reframing Research Ethics in Engineering Education Research To Improve Research Quality. *Journal of Engineering Education*, *107*(3), 362–379. https://doi.org/10.1002/jee.20222
States, N. L. (2013). *Next Generation Science Standards: For States, By States*. Washington, DC: The National Academies Press.
Thiry, H., Weston, T., Laursen, S., & Hunter, A. (2012). The benefits of multi-year research experiences: differences in novice and experienced students' reported gains from undergraduate research. *CBE Life Sciences Education*, *11*(3), 260–272. https://doi.org/10.1187/cbe.11-11-0098
Thiry, Heather, Laursen, S. L., & Hunter, A. (2005). What Experiences Help Students Become Scientists ? A Comparative Study of Research and Other Sources of Personal and Professional Gains for STEM Undergraduates, *82*(4).
Toulmin, S. E. (2003). *The Uses of Argument, Updated Edition* (2nd ed.). New York, NY: Cambridge University Press.
Walker, J. D., Cotner, S. H., Baepler, P. M., & Decker, M. D. (2008). A Delicate Balance: Integrating Active Learning into a Large Lecture Course. *CBE-Life Sciences Education*, *7*, 361–367.
Walther, J., Sochacka, N. W., & Kellam, N. N. (2013). Quality in interpretive engineering education research: Reflections on an example study. *Journal of Engineering Education*, *102*(4), 626–659. https://doi.org/10.1002/jee.20029
Watkins, J., & Elby, A. (2013). Context dependence of students' views about the role of equations in understanding biology. *CBE Life Sciences Education*, *12*(2), 274–286. https://doi.org/10.1187/cbe.12-11-0185
Webster, D. M., & Kruglanski, A. W. (1994). Individual Differences in Need for Cognitive Closure Individual Differences in Need for Cognitive Closure. *Journal of Personality and Social Psychology*, *67*(6), 1049–1062. https://doi.org/10.1037/0022-

3514.67.6.1049
Weinstock, M., Kienhues, D., Feucht, F. C., & Ryan, M. (2017). Informed Reflexivity: Enacting Epistemic Virtue. *Educational Psychologist*, *52*(4), 284–298. https://doi.org/10.1080/00461520.2017.1349662
Whewell, W. (1847). *The Philosophy of the Inductive Sciences: Founded Upon Their History, Volume 2*. London, UK: J. W. Parker.
Yin, R. K. (2018). *Case Study Research and Applications: Design and Methods* (6th ed.). Los Angeles, CA: SAGE Publications.