Title

What Makes a Scientific Research Question Worth Investigating? Students' Epistemic Criteria and Considerations of Contribution

Author

Berson, Eric

Publication Date

2012

What Makes a Scientific Research Question Worth Investigating? Students' Epistemic Criteria and Considerations of Contribution

by

Eric Bruckner Berson

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Education

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge:

Professor Kathleen E. Metz (Chair)
Professor Barbara Y. White
Professor Miguel Altieri

Fall 2012

What Makes a Scientific Research Question Worth Investigating?
Students' Epistemic Criteria and Considerations of Contribution

by

Eric Bruckner Berson

Abstract

What Makes a Scientific Research Question Worth Investigating?
Students' Epistemic Criteria and Considerations of Contribution

by

Eric Bruckner Berson

Doctor of Philosophy in Education

University of California, Berkeley

Professor Kathleen E. Metz, Chair

This dissertation introduces the construct of *worthwhileness* as an important aspect of students' *practical* epistemologies of science (Sandoval, 2005). Specifically, it examines how students conceptualize what makes a scientific research question worthwhile, through a close analysis of the criteria they use for evaluating scientific research questions. Elementary (n=21) and high school students (n=21) participated in this study. As part of semi-structured interviews, students engaged in three novel tasks designed to elicit the epistemic criteria they use to evaluate scientific research questions in a variety of contexts.

Findings indicate that elementary and high school students alike could engage in the practice of evaluating the worth of scientific questions. The criteria they employed included degree of interest, difficulty, and the contribution of questions to knowledge or to solving a problem. The criteria students considered varied by context. Several key differences emerged between the reasoning of the two grade cohorts. High school students tended to place more weight on the contribution of the research question. Also, the criteria reflected in the high school students' judgments of the scientific value of individual questions more closely accorded with the criteria they identified retrospectively as the basis of their judgments. Furthermore, the older cohort more often rationalized the selection and sequence of research questions within a single domain on the basis of epistemic contingency between questions.

How students conceptualize what makes a scientific research question worthwhile constitutes a key aspect of students' epistemic reasoning. It is particularly important to understand how students judge the worthwhilness of scientific research questions given the central epistemic role of research questions in scientific inquiry.

Table of Contents

List of Tables

List of Figures

Acknowledgements

Writing this dissertation has been a journey that I could not have completed without substantial support and guidance from others. It has been a process of learning and growth and I have many people to thank.

First, I am deeply grateful to my advisor and mentor, Dr. Kathleen Metz, for guiding me in my graduate studies and dissertation research. The passion and rigor that she brings to her work has been an inspiration to me. She has pushed my thinking in so many ways and has provided me with continued guidance and support as I conducted my dissertation research. In working with her, I have become a better thinker and have learned how to conduct educational research. I am so thankful for the immense amount of time she has invested in me and I am honored to have her as my mentor.

I want to thank Dr. Barbara White for inspiring me with her important work and contributions to the field. Her perspective and feedback on my dissertation has been invaluable. I always leave our conversations feeling excited and motivated about conducting research.

Special thanks to Dr. Maryl Gearhart who, from day one, has been a source of constant support throughout my time in graduate school. She always made herself available to listen to whatever struggle I was facing or decision I had to make, offering heartfelt advice and looking out for my best interests even as my interests and goals evolved over time. Her support and encouragement helped get me through graduate school. I could not have made it through without her support.

I want to thank Dr. Mark Wilson for the research opportunities he provided me during my time in graduate school. I have learned so much from his courses and from working with him on a variety of fascinating projects. I appreciate his support, advice and encouragement. Thank you to Dr. Miguel Altieri for exposing me to agroecology, a field that inspired my research, and for serving as a member of my committees. Thank you to Albie Miles and Houston Wilson for their expertise and valuable consultations on my dissertation work. Thank you to Jiwandeep Kohli and Savio Chan for their assistance on my dissertation.

At the Graduate School of Education, I have been fortunate to be part of a community of fellow graduate students and recent graduates who have been a source of support throughout my time in school. Thank you to my colleagues and friends: Tia Madkins, Lina Haldar, Niral Shah, Jennifer King Chen, Dr. Uyen Ly, Dr. Nicole Wong and Dr. Stephanie Sisk-Hilton.

My family has been so amazingly supportive of me as a graduate student and while writing my dissertation. Thank you to my nephew Noah and niece Daisy who were my first two subjects as I piloted my dissertation interview tasks. A special thanks to my parents for their endless encouragement during my entire life and for taking so much interest in my research.

I want to express how deeply appreciative I am for the support of my wife Ariel and son Jonah. Ariel moved with me to Berkeley over five years ago so that I could attend graduate school. She has been patient and supportive of me every single day since. I could never have completed the program without her endless love and care. Thank to my son Jonah, who was born shortly after I completed my oral exam, and will be two years old this month. His is an absolute blessing and his curiosity and excitement about the world inspires me to continue working in the field of science education research.

Lastly, I want to express my gratitude to the students who participated in my dissertation study. They volunteered their time and shared their thinking with me. Without their openness, this dissertation research would not have been possible.

Chapter 1: Introduction

Scientific inquiry is the engine that drives the expansion of scientific knowledge. Scientists ask questions about the world and develop theories and models by conducting investigations, collecting data and constructing new understandings about how the world works. The role of scientific questions has also been a central focus of science education efforts. The National Research Council's new *Framework for K-12 Science Education* highlights the asking of scientific questions as one of eight focal scientific practices in which students should engage. Specifically, the standards expect students to formulate their own scientific questions about the world, refine questions such that they can be empirically investigated and distinguish between scientific and nonscientific questions (NRC, 2012).

While scientific questions can facilitate the construction of scientific knowledge, not all scientific questions are necessarily of equal epistemic value. In other words, some scientific questions may do more to advance scientific understanding than other questions. If asking scientific questions is a core practice of science and of central importance for students who are learning to engage in scientific practice, then it is crucial that students come to understand what makes some scientific questions more powerful than others.

So what "counts" as a good scientific question? What makes one scientific question more worthwhile to investigate than another? Thinking about the merit of scientific questions and their utility for both scientific inquiry and constructing scientific knowledge is of fundamental epistemic concern. Students' judgments about what makes a scientific question worthwhile hinge on their beliefs about how science works and the role of scientific questions in the enterprise of science. Thus, the formulation and evaluation of scientific questions are intertwined with one's understanding of the nature of science and the epistemic criteria that one holds about what makes a scientific question a worthwhile one.

This dissertation closely examines the epistemic beliefs about the nature of science that students hold as reflected in their judgments about which scientific questions are worthwhile to investigate. The dissertation's focus on epistemic criteria builds on an emerging line of research in science education that is specifically concerned with students' beliefs about the nature of science and their views of particular epistemic forms such as models, explanations and evidence. I focus on the epistemic criteria that students use as they engage in an epistemic practice of deciding which questions are worthwhile to investigate.

In this first chapter, I make an argument for why this research focus is strategic by reviewing relevant research literature. I begin with a brief elaboration of the goals of science education relating to students' understanding of the nature of science and the reasons why understanding the nature of science is so important for students. Next, I outline a trajectory in the literature that has led to the current focus on students' epistemic criteria. I then review the research literature about scientific questions in science education including the benefits and challenges of students formulating their own scientific research questions. I conclude with an argument for "worthwhileness" as a strategic construct that merits attention in science education and present the core research questions that frame my dissertation research.

Nature of science and the goals of science education

Science is both a body of knowledge and a way of knowing. Therefore, it is important for students to understand the nature of scientific knowledge and how scientific knowledge is constructed. To emphasize the importance of students' understanding of the nature of science (NOS), the National Research Council elevated NOS as one of four central strands of science proficiency in their seminal publication, *Taking Science to School* (NRC, 2007). The recently released Framework for K-12 Science Education also specifically stresses the importance of reflecting on scientific practices and how they contribute to the construction of scientific knowledge (NRC, 2012). Developing students' understanding of how science works has long been a priority for science educators (Rudolph, 2005).

There are many arguments for why it is important for students to understand the nature of science including ones made on economic, utilitarian, cultural and moral grounds (Driver, Leach, Millar & Scott, 1996). One prominent argument is that understanding how scientific knowledge is constructed is important for one's productive participation in democracy (Roth & Desautels, 2002; Kolsto, 2001; Driver et al., 1996). Citizens need to understand how science works and the basis of scientific claims in order to effectively participate in dialogue concerning pressing issues involving science (e.g., global warming, genetic engineering etc.)

Another argument for the importance of teaching NOS is that by understanding the nature of science and how scientific knowledge is constructed, students will be better positioned to engage in scientific practice themselves (Sandoval, 2005). As Driver et al. (1996) note, there is a risk that students "may not appreciate that the aim of science is to establish explanations for the behavior of natural objects and phenomena which can command widespread acceptance. Some may simply see science as the accumulation, by observation and measurement, of 'facts' about the natural world."

Epistemic cognition

Philosophers have long pondered epistemological issues of the nature of knowledge, truth and how we know what we know. Within the field of psychology, research on epistemic cognition has investigated how individuals think about the nature of knowledge. Kitchner (1983) explains that whereas cognition refers to such tasks as problem solving, and metacognition refers to the process of self-monitoring one's own cognition, *epistemic cognition* includes "the individual's knowledge about the limits of knowing…the certainty of knowing and the criteria for knowing" (p. 225). Kitchener argues that epistemic cognition is particularly important for reasoning about ill-structured problems that do not necessarily have one right answer. If people believe that knowledge is absolute and known by someone, then they may be likely to think that there is only one way to solve the problem, thereby constraining their potential solutions.

A pioneer in the field of epistemic cognition, psychologist William G. Perry conducted a seminal study of the intellectual and cognitive development of college students (1970). He interviewed students by asking them to reflect on their own educational experiences. From these interviews, Perry proposed a developmental scheme with levels that characterize students' different views of the source of truth, authority and one's own role in knowledge construction (1970). Since then, a number of researchers have posited different models of the development of individuals' epistemic beliefs (see Hofer & Pintrich, 1997 for a review). In Hofer & Pintrich's (1997) view, an individual's collective epistemic beliefs are best viewed as a theory of knowledge or a *personal epistemology.* They proposed a framework for characterizing individuals' personal epistemologies focused on two core epistemic beliefs: beliefs about the nature of knowledge and the nature or process of knowing. Within these two general constructs are individuals' beliefs about the certainty of knowledge, the simplicity of knowledge, the source of knowledge and justifications for knowing. This framework has informed subsequent research in students' personal epistemologies of science.

Developing students' epistemologies of science

If a major goal of science education is to develop students' personal epistemologies of science, it is important to consider what should be the target model of the nature of science for students to learn. Although there are differences among philosophers of science about what constitutes the nature of science (Alters, 1997), there have been several attempts by researchers in science education to distill core principles of the nature of science for the purpose of instruction (Lederman, Abd-El-Khalick, Bell, & Schwartz, 2002; McComas & Olson, 1998; Osborne, Collins, Ratcliffe, Millar, & Duschl, 2003; Sandoval, 2005). Sandoval (2005) proposed a consensus list of four broad epistemological themes that include the view that scientific knowledge is constructed, scientific methods are diverse, scientific knowledge comes in different forms and scientific knowledge can vary in certainty.

The notion that scientific knowledge is constructed by scientists rather than simply residing in the world to be discovered has been a central focus of several important efforts to characterize students' epistemic beliefs. Carey, Evans, Honda, Jay, & Unger (1989) and Driver et al. (1996) have each proposed frameworks derived from empirical research findings about how students think about the construction of scientific knowledge. Carey et al. (1989) and Carey & Smith (1993) put forth a three-tiered framework that reflects increasing sophistication in student understanding of the role of theory in scientific experimentation. In level 1, students do not differentiate between scientific ideas and the experiments used to test them. In other words, science is "when you do an experiment and see what happens." In level 2, students make an explicit distinction between ideas and the means to test them. In level 3, students view the goal of science as the construction of deeper explanatory theories of the natural world.

Driver et al. (1996) also proposed a three-tiered framework for analyzing students' epistemologies of science. In the Driver framework, the levels describe students' epistemologies as phenomenon-based, relation-based and model-based. A phenomenon-based view of science means that students think that science consists of making direct observations about the world, without distinguishing between description and explanation. A relation-based epistemology views explanations as a form of empirical generalization about relationships between observable variables that exist in the world. Students with a model-based epistemology view science as the evaluation of models using empirical evidence with a clear distinction between description and explanation. In a cross-sectional study, the authors observed increasingly sophisticated epistemologies in students at higher grades, but even high school students did not commonly surpass a relation-based understanding of science.

Instructional models for teaching about the nature of science

The challenge for science educators and researchers is to design instruction that supports students in developing a more sophisticated understanding of the nature of science. In his review of the field, Lederman (2007) concludes that the many efforts and approaches to improving student understanding on measures of NOS have been met with limited success.

Historical approaches to teaching NOS are predicated on the notion that incorporating the history of science in science instruction will improve students' understanding of NOS. However, these approaches have had limited success (Solomon, Duveen, Scot & McCarthy, 1992). In contrast, implicit instructional approaches reason that students will develop an understanding of NOS by virtue of engaging in hands-on scientific inquiry activities. However, implicit approaches to teaching NOS have also failed to show significant gains on assessments of the nature of science (see Khishfe & Abd-El-Khalick, 2002).

Despite the difficulties in improving students understanding of NOS, one instructional approach for developing students' epistemologies of science has shown particular promise. Several researchers have explicitly emphasized student reflection about scientific practice as a central feature of the instructional model (Smith et al., 2000). Khishfe & Abd-El-Khalik (2002) conducted a controlled study to determine the impact of reflective discussion about NOS on sixth graders' understanding of the tentative, empirical, inferential, imaginative and creative aspects of the nature of science. Both cohorts participated in scientific inquiry activities but only one cohort had reflective discussions about how they constructed scientific knowledge during the inquiry activities. The authors found that students who engaged in explicit reflection about NOS performed better than the control group on measures of NOS. Khishfe & Abd-El-Khalik (2002) argue that "treating NOS as a cognitive instructional outcome, with the consequence that specific instructional activities are geared toward teaching such an outcome" can increase students understanding of the nature of science.

In another study, Carey et al (1989) examined changes in students' epistemologies of science after they participated in an instructional unit on yeast. Understanding the nature of science constituted an explicit focus of the unit. Instruction provided students with structured opportunities to reflect on their scientific practices and to discuss how scientific knowledge was being constructed in the course of the inquiry project. Carey et al. (1989) found modest gains in students' epistemological understanding following their participation in the unit.

Schwarz & White (2005) also developed an instructional model focused on the nature of scientific models and the role they play in science. The authors referred to this understanding as "meta-knowledge" about modeling. In the controlled study, students in the cohort with a "meta-knowledge" layer with explicit opportunities for reflection about models outperformed the control group on measures of understanding of models and the nature of science.

These studies recognize the importance of treating the nature of science as an explicit cognitive and instructional goal. Furthermore, the pedagogical strategy of making the nature of science explicit to students has been effective at developing students' epistemologies of science. As discussed in the next section, highlighting and reflecting on epistemic criteria are strategic approaches for making epistemic thinking explicit in the science classroom.

Epistemic forms and epistemic criteria

The power of explicitly reflecting on how scientific practices function to construct scientific knowledge is also a strategic reason to highlight the particular *epistemic forms* that constitute the enterprise of science (Collins & Ferguson, 1993). Collins & Ferguson (1993) define epistemic forms as "target structures that guide scientific inquiry." Building on Collins & Ferguson's construct, generalized forms of scientific arguments, explanations and models can also be conceptualized as epistemic forms given their function for constructing scientific knowledge.

If students are to understand how science works, it is important for them to understand the different epistemic forms that are used to construct scientific knowledge (NRC, 2012). Furthermore, it is crucial for students to have a deep understanding of what characterizes a particular epistemic form in science and to be able to distinguish between high and low quality (i.e., knowing the difference between a good and not-as-good scientific explanation). The ways that students apply epistemic criteria are direct reflections of students' epistemology of science. In reference to epistemic criteria for scientific arguments, Duschl & Osborne (2002) explain, "it is the role of science education, therefore, to explore the criteria by which such evaluations are made, and explain how those criteria are themselves justified" (p. 45).

One line of epistemic cognition research has examined how students develop epistemic criteria for particular epistemic forms in science. Pluta, Chinn, & Duncan (2011) note that "understanding criteria and criteria-related practices is an important part of learning how to participate in science, as well as understand the nature of science (NOS)…the use of criteria is embedded in the modeling, argumentation, and evidence evaluation practices that are central to inquiry curricula"(p.487). They also argue that debating and applying epistemic criteria can support classroom learning communities around social scientific practice.

Prior research has examined students' use of epistemic criteria in evaluating various epistemic forms. Samarapungavan (1992) studied students' epistemic criteria for evaluating scientific theories. Duschl's (2007) work focused on epistemic criteria for argumentation. Hogan & Maglienti (2001) compared students' criteria for scientific conclusions with those given by lay adults and scientists. Smith et al. (2000) studied students' criteria for scientific beliefs. Research has also focused on scaffolds for developing epistemic criteria for scientific explanations (Sandoval, 2003; Sandoval & Reiser, 2004). Scientific models have also been a major focus for research on epistemic criteria for the goodness of scientific models (Penner, Giles, Lehrer, & Schauble, 1997; Pluta et al., 2011; Schwarz & White, 2005). While these prior studies have investigated a range of epistemic forms in science, there has been surprisingly little research focused on epistemic criteria as applied to scientific questions.

Scientific questions in science and science education

Scientific questions anchor scientific inquiry. Asking scientific questions is a hallmark practice of authentic science (Chinn & Malhotra, 2002). Scientific questions can emerge from puzzling observations that cause cognitive dissonance. Experiencing phenomena first-hand stimulates students' curiosity which can, in turn, motivate inquiry (Simon, 2001). Scientists also may have theoretical questions about how the world works that they then test through experimentation. Whether questions are sparked by observation or lead to subsequent observation or both, scientific questions function to propel and direct scientific enterprise.

Scientific questions also take many different forms and play many different roles in the teaching and learning of science. Chin & Osborne (2008) conducted a review of research on scientific questions in science education. They identified the diverse roles that scientific questions can play in science learning including facilitating direct knowledge construction, discourse, self-monitoring and increased motivation of students. Chin & Osborne (2008) also outlined the role of scientific questions in the teaching of science including for formative assessment, evaluating student thinking, stimulating inquiry and provoking critical reflection. These authors also found that scientific questions stimulate scientific argumentation between students (Chin & Osborne, 2010).

Scientific research questions are questions that motivate scientific inquiry. These include broad questions that frame long-term research agendas or specific empirical questions investigated through individual studies or experiments. It is by investigating scientific research questions that scientists generate evidence that contributes to the formation of theoretical explanations of how the world works. By anchoring scientific investigations, scientific research questions drive associated inquiry practices such data collection, analysis, inference, model-building, explanation-building and argumentation. Scientific research questions have a particularly important role in the construction of scientific knowledge and are fundamental to viewing science as a way of knowing. Therefore, it is crucial for students to learn how to formulate and pursue scientific research questions of their own.

Student-generated scientific research questions

Being able to develop and pursue a researchable question is emphasized as a core scientific practice in the national *Framework for K-12 Science Education* (NRC, 2012). There are several cognitive affordances for students in formulating their own research questions. First, when a student asks a scientific question, the question emerges from a gap between students' prior knowledge and new information or stimuli that the student encounters (Chin & Osborne, 2008). Students' questions reflect an attempt to gain information that can extend their understanding at a particular time when their conceptual structures are activated. As Chouinard (2007) explains, "asking such questions is a mechanism that can fill in gaps in the child's knowledge, help resolve internal disequilibrium, and – importantly – guide the direction of the child's thought, precisely when the child is trying to resolve such equilibrium"(p.3). Thus, questions are an important mechanism for students in constructing new knowledge.

Second, the formulation of questions requires the important cognitive processes associated with navigating a problem space (Newell & Simon, 1972). In science, this means taking a particular problem or puzzling phenomena that is "ill-structured" and determining how to frame the problem in a way that can be investigated (Simon, 2001). When students generate their own research questions, the questions have the potential of being within the students' own zone of proximal development (Vygotsky, 1978). When the question is appropriately challenging for the student, the student is primed for learning because the challenge is just beyond what they can achieve by themselves, thus in their zone of proximal development.

There are also motivational affordances of encouraging students to develop their own research questions. Chin and Kayalvizhi (2005) reported that kids were more enthusiastic about investigating questions that they had generated themselves, compared to the investigative questions that were in their activity books. When students generate their own questions, they feel a sense of ownership. Students' motivation to pursue their own questions can also influence their cognitive development. Pintrich et al. (1993) cite several motivational behaviors including students' goal orientations, interests, value beliefs, self-efficacy beliefs and control beliefs that can influence students' cognitive development. For example, if a student has an intrinsic goal orientation, he is more likely to persist in working on a problem, thereby avoiding closure while seeking a specific answer (Kruglanski, 1990). In contrast, students who are motivated by grades or other extrinsic factors may be more inclined to settle for any answer and move on.

Furthermore, Katz & Assor (2007) argue that choice (i.e. selecting a research question) can lead to intrinsic motivation if certain requirements are met. Based in the self-determination theory of motivation (Deci & Ryan, 2000), choice is motivational when the choice fulfills students' psychological need for autonomy, competence and relatedness. Thus, allowing students the opportunity to choose a question that interests them can result in students' intrinsic motivation and commitment to that question. This helps propel students in their pursuit of the answer to their question.

Despite the potential benefits of students generating their own research questions, students are rarely afforded the opportunity to do so in typical science classrooms. In their analysis of high school biology laboratory manuals, Germann, Haskins & Auls (1996) found that the curricula they reviewed seldom required students to pose a question to be investigated, formulate hypotheses, design procedures or formulate new questions based on findings from completed investigations. Chinn & Malhotra (2002) reviewed published science curriculum *and* researcher-developed inquiry curriculum to compare the characteristics of the inquiry tasks with those of authentic inquiry. None of the textbook inquiry tasks and only 12% of the researcher-developed tasks asked students to generate their own research questions.

When students are afforded the opportunity to generate their own questions for empirical investigation, what is the inspiration for students' questions? Students' interests in topics for investigation may simply come from passing thoughts, fleeting concerns, phobias, obsessions, or fascination with media-related characters (Katz & Chard, 1998). Chin & Chia (2004) found that the sources of inspiration for students' questions included curiosity born from personal encounters, family concerns, wonderment about advertising messages, cultural beliefs, and issues that arose from prior lessons in school. These findings indicate that there is a wide range of sources of inspiration for student questions. While this finding could be interpreted as promising in that students are looking to science to answer questions from daily life, the finding also raises some important concerns.

One major concern about providing students the opportunity to investigate their own questions is that the questions students develop may be conceptually impoverished. When teachers determine the question, they have more control over the direction of inquiry and can target specific scientific concepts. In her study of open-inquiry with 6th graders, Keys (1998) found that "allowing students to generate their own questions compelled the teachers to leave scientific concept development open-ended … specific content learning from these experiences was idiosyncratic to each group of students" (p. 313). Blumenfeld (1991) and Metz (2011) further elaborate this concern:

> *"Balancing students' need for choice and control in the selection of problem questions, approaches, and artifacts so that they feel "ownership" with the need to have students address and learn content defined by curricular mandates and requirements poses a significant dilemma."* (Blumenfeld et al., 1991, p. 377)
>
> *"There is a trade-off between student regulation over the inquiry and its conceptual power. Curriculum writers can craft inquiry that more powerfully address underlying ideas or theories than young -- or older- students are likely to develop on their own. However there is a significant benefit in having students assume considerable responsibility for structuring their own investigation, including their investment in the enterprise and deeper mastery of the practices of inquiry."* (Metz, 2011, p.17)

Design-based researchers have developed instructional models that successfully scaffold the full cycle of student-directed inquiry (e.g. Roth & Bowen, 1993; Zion et al., 2004). In these design studies, students' inquiry is structured within the context of a particular conceptual frame such as building models of pond ecologies (Lehrer & Schauble, 2008), studying cricket behavior (Metz, 2004) or designing and testing computer models of Newtonian physics (White & Frederiksen, 1998.) The specific conceptual frames of these studies position students to ask more conceptually rich questions, but consequently constrain the space of potential research questions that students can ask.

Chin & Kayalvizhi (2002) explicitly investigated the process of question generation and formulation in inquiry contexts. Students were asked to generate questions for scientific inquiry in a setting that might be more similar to a science fair in which any topic is eligible. The study found that only 11.7% of the questions 6th graders generated were amenable to empirical investigation (i.e. "investigable"). These findings echo prior research that highlights another major concern about student-generated research questions: student-generated questions may not be amenable to empirical investigation.

Taken as a whole, the research on student-generated scientific research questions raises some important concerns about the quality of the questions that students generate. Drawing on the epistemic cognition literature, it is possible that supporting students' understanding of the epistemic function of research questions in science could be a strategic avenue for teaching students how to develop higher quality questions through a more sophisticated understanding of the nature of science. From this perspective, it is important to develop students' personal epistemologies of science through an explicit focus on the epistemic criteria for what makes a scientific research question a good one. In the next section, I review extant research on efforts to teach students about criteria for good scientific research questions.

Teaching students criteria for good scientific research questions

The most common approach to improving the quality of student-generated research questions has been to expose students to instruction focused on specific attributes of good questions. Several studies have shown that, through relatively short instructional interventions, students can learn how to "operationalize" their questions (Alfke, 1974; Allison & Shrigley, 1986) or make their questions more "researchable" (Cuccio-Schirripa & Steiner, 2000; Hofstein, Navon, Kipnis, Mamlok-Naaman, 2005). However, these studies have almost exclusively focused on the empirical feasibility of the question.

Lucas, Broderick, Lehrer, & Bohanan (2005) developed criteria for good questions in conjunction with the students in the study. A good research question was answerable, worded in a way that can be researched, crafted with a hypothesis in mind and genuine, a question to which the student does not already know the answer. Chin & Kayalvizhi (2002) propose a similar list of criteria for good student questions. The questions must be doable, interesting, practical and unknown to the students.

White and colleagues (2005, 2009, 2011) argue that developing students' "meta-questioning knowledge" is a crucial aspect of students' understanding of scientific inquiry. They describe meta-questioning knowledge as knowledge about different types of research questions and how investigating those question types relates to developing other epistemic forms, such as structural, causal and process models, that embody a scientific theory for any given domain. To develop students' meta-questioning knowledge, White et al. created a software-based epistemic scaffolding program called *Inquiry Island*. The software included various "advisors" that served as coaches for different phases of the inquiry cycle.

One software advisor, Quentin Questioner, guided students to consider specific criteria for a good scientific research question (White & Frederiksen, 2005; White et al., 2011). Quentin considered a good question to be one that is built on prior research, interesting, important and worthwhile, open to alternative hypotheses based on different people's theories and possible to investigate given available resources. He characterized each goal, provided motives for why each goal was important and strategies for achieving that goal. Then, when students had generated a potential research question, he had students evaluate their question against each goal. If the students gave themselves low ratings for any of the goals, he offered specific advice (which included strategies, questions to ask themselves, and examples) for how to achieve that goal and thereby create a better research question. A particularly important feature of the White et al. approach was that it encouraged students to develop theories before generating questions. This feature was strategic for scaffolding the theoretical basis for a research question.

While many of these studies involved the teaching of specific criteria to students or advising how to make a question more "researchable," they did not examine resulting changes in students' own epistemic beliefs. A different type of research is needed to understand the extent to which students *understand* what makes a question a good one and the basis upon which they make judgments about scientific questions.

Students' epistemic criteria for good scientific research questions

To better understand how students' think about what motivates scientists to ask scientific questions, Driver et al. (1996) and Carey et al. (1989) interviewed students directly. The Driver et al. (1996) study interviewed 9, 12 and 16 year old students about a wide range of epistemological issues, one of which was their view of what distinguishes a scientific question from a non-scientific question. The interviewers told students that scientific questions are questions "that scientists might want to find out more about." The authors gave this vague definition so as to allow students to more fully interpret the basis on which scientists might make that decision. The authors then shared a set of eleven questions with the students and asked them to indicate if the question was "scientific or not" and to explain the basis for their judgment.

There were several important findings from the Driver et al. (1996) study. Students' rationales tended to focus on three features: empirical "investigability," the nature of the domain of the question and the perceived personal and professional characteristics of scientific work. Younger students thought that scientific question could be answered simply by looking, whereas older students tended to view questions as addressing underlying mechanisms or theories. Students who considered a question to have broad social relevance tended to cite that as a reason for why scientists would investigate that question. While younger students more often considered the personal interests of the scientist, older students were more likely to consider the professional interests of the scientists. While the Driver et al. (1996) study was effective at interpreting students' criteria for what makes a question scientific, it did not specifically investigate students' epistemic criteria for what makes a scientific question a particularly good one.

Carey et al. (1989) used a different approach to investigate students' beliefs about the goals of science. Rather than eliciting criteria about specific scientific questions, the authors interviewed students about the goals of science and scientists' motivations for experimentation. The authors asked 7th grade students who had participated in a three-week science curriculum questions such as: What do you think the goal of science is? Where does a scientist get a hypothesis? How does a scientists decide what experiments to do? By analyzing students' responses to these questions, Carey et al. (1989) found that, before instruction, students tended to think that scientists are motivated by their own whims and desires rather than seeking particular information. Students did not consider any particular purpose for the experiments other than to see what happens. Following instruction, students were more likely to understand that scientists experiment in order to test a particular idea or question. The most sophisticated students explained that scientists build on their prior knowledge and experience or test and develop new ideas through experimentation.

While the Driver et al. and Carey et al. studies begin to investigate students' ideas about the role of scientific questions, they do not present a complete picture of students' epistemic criteria as applied to scientific research questions. The Driver et al. study provides students with the criterion that they should use in the task (scientific vs. not scientific.) While this criterion is open to interpretation, the demands of the task limit the full range of criteria that students might apply to evaluating scientific question in other contexts. Furthermore, deciding if a question is scientific or not does not engage students in thinking about the particular epistemic utility of the question. While two questions may both be scientific ones, they may not do equal "work" in terms of advancing the goals of science. While the Carey et al. study more explicitly focuses on issues of epistemic utility, the interview protocol does not require students to make epistemic judgments about particular scientific questions and therefore does not elicit the criteria that students apply specifically to scientific questions.

Overall, research on epistemic criteria of scientific questions is limited. The research has not explicitly investigated the extent to which students even consider the epistemic utility of a scientific research question or the relative epistemic power of different scientific questions.

‘Worthwhile’ scientific research questions

Scientists pursue scientific research questions that they consider to be worthy of pursuit. Scientific journals and grant funding agencies decide to publish and fund worthwhile research. Scientists depend on funding to support their research programs. In their proposals, scientists must make a convincing case that their research questions are worthwhile. Sternberg’s guidelines for proposal writing in the social sciences illustrate the central requirements in making a case for worthwhile scientific research (2003, p. 236-237):

1) Clearly state the “big question” you hope to address.
2) Show why the big question is important.
3) State how your work builds on and departs from work that has been done before.
4) State your theory and how it relates to the theories of others.
5) Show why your theory is better than its competitors.

Worthwhile research addresses an important question. Specifically, a worthwhile question “builds on and departs from” prior work in the field. In other words, a worthwhile scientific research question makes a contribution of some kind. But what is the nature of the contribution that a scientific research question needs to make for it to be deemed worthwhile?

In his seminal book on scientific research, *Pasteur’s Quadrant*, Donald E. Stokes (1997) traces how the goal structures of scientific research have been conceptualized. He explains that the goal or contribution of scientific research has been viewed either in terms of a quest for understanding (basic research) or with consideration of use (applied research).

Stokes argues that scientific research need not be dichotomized. Rather, scientific research can contribute simultaneously to understanding *and* use. He calls this “use-inspired basic research” and refers to that goal of research as “Pasteur’s Quadrant” in deference to the major basic and applied contributions of Louis Pasteur (Figure 1). Stokes suggests that decisions to proceed with research should entail a judgment about both scientific promise and social value. Thus, Stokes’ frame views the contribution of a worthwhile research question both in terms of general scientific understanding and a particular application.

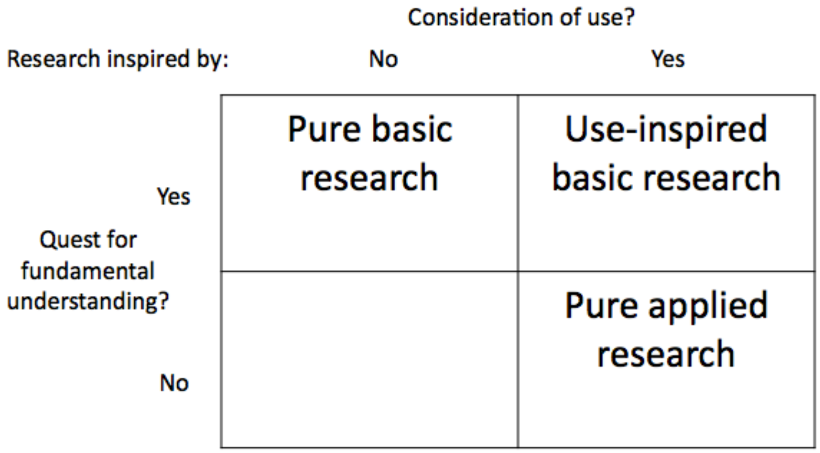


Figure 1: Quadrant Model of Scientific Research (Stokes, 1997, p. 73)

Students' conceptualizations of 'worthwhileness' as a strategic construct

Given the importance of scientific questions to scientific inquiry and the extent to which students' epistemic criteria reflect their tacit beliefs about how science works, encouraging students to consider what makes a scientific research question worthwhile to investigate is a strategic avenue of research. Chinn, Buckland and Samarapungavan (2011) propose an expanded framework for epistemic cognition, in which they advocate for research addressing students' *epistemic values,* a feature of their framework for epistemic beliefs. Chinn et al. explain that "people adopt various epistemic aims because the resulting epistemic achievements have value to them" (p. 148).

The construct of "worthwhileness" builds on Chinn et al.'s notion of epistemic value, applying the idea to scientific research questions in particular. Students' judgments about worthwhile scientific questions are informed by their epistemic values and the particular value they attribute to specific questions. Students' considerations of epistemic value factor into and are reflected in the epistemic criteria they use to evaluate the questions. While prior research has examined epistemic criteria in relation to other epistemic forms, no research has fully examined students' epistemic beliefs about worthwhile scientific research questions.

I argue that the construct of 'worthwhileness' is strategic because considerations of worthwhileness are evoked in the practice of formulating and critiquing scientific questions. A decision to pursue a research question hinges on a determination that the research question will make a worthwhile contribution of some kind. While there are certainly other considerations that may play a part in the decision (i.e., cost, difficulty, etc.) the contribution of the question is paramount.

Examining how students conceptualize what makes a scientific research question worthwhile contributes to our understanding of students' epistemologies of science. Sandoval (2005) argues that most research on student understanding of the nature of science has focused on assessing students' formal epistemology, or their beliefs about professional or formal science. Sandoval (2005) suggests that future research should also focus on students' *practical epistemologies*. Students' practical epistemologies are reflected in the entailments of students' epistemologies of science in the context of decisions students make while engaged in a process of scientific inquiry. Rather than focusing on students' understanding of formal science, a practical epistemological approach examines the epistemological basis of student decision-making about what counts as evidence, claims, etc. Therefore, it is strategic to engage students in practices of science during which they articulate how they are making decisions. The basis of the decisions can be analyzed for evidence of students' epistemic criteria.

Dissertation research

In this dissertation study, I investigate how students reason about what makes a scientific research question worthwhile. In a series of interview tasks, I invite students to play the epistemic game of critiquing scientific research questions, thereby eliciting students' epistemic criteria in practice. I examine what criteria students use in the rationales for their decisions, how they apply the criteria, and differences in reasoning by grade cohort.

I also analyze student epistemic reasoning in multiple contexts. I compare the criteria students use as they are evaluating scientific questions and also how they retrospectively reflect on the criteria they used. This analysis is important for examining the implicit versus explicit use of epistemic criteria by students. I also compare how students reason about what makes a question worthwhile for professional scientists to investigate versus criteria that students think should be applied to scientific research questions in a school science fair context. If students have different images of school and professional science, then it is possible that the epistemic criteria they apply in each context may be different as well.

Finally, I examine how students evaluate a set of related research questions on the same topic to examine finer-grain distinctions between the relative utility of different questions. This analysis also examines how students coordinate multiple research questions as part of a plan to investigate a particular idea. The coordination of questions invites students to consider issues of contingency among research questions as part of a research program.

The dissertation research makes a contribution to the research literature on students' epistemic criteria broadly and, specifically, in relation to scientific research questions. This research also has instructional implications for possible ways of teaching science in schools. Findings from this research can contribute to pedagogical methods and epistemic scaffolds for supporting students in formulating more sophisticated scientific research questions while also developing their epistemologies of science.

The analysis in this dissertation is organized around three primary research questions:

- Chapter 3: On what basis do students form judgments about what makes a scientific research question worthwhile in different contexts?

- Chapter 4: How do students conceptualize the particular contribution of scientific research questions?

- Chapter 5: How do students evaluate and coordinate a set of scientific research questions to investigate an idea?

Before presenting the analysis and empirical findings, I begin with a description of the design, subjects, and methodology of the dissertation study. Following the analytical chapters, Chapter 6 includes a discussion of the contribution of the dissertation research, implications of the study and suggestions for future research.

Chapter 2: Methods

This dissertation seeks to understand how students conceptualize what makes a scientific research question worthwhile. Conceptions of *worthwhileness* are embedded in students' personal epistemologies of science. As Sandoval (2005) argues, it is important to consider students' epistemologies of science in the context of their inquiry practice. In other words, students' epistemic beliefs are reflected in the practical decisions that students make in the course of engaging in scientific practice. Sandoval (2005) refers to these tacit beliefs as students' *practical epistemologies.*

To examine students' practical epistemic beliefs in relation to scientific research questions, this study invited students to engage in the epistemic game of evaluating scientific research questions and determining which questions the students judged to be worthwhile for empirical pursuit. Evaluating scientific research questions is an authentic scientific practice, and it is in making judgments about the questions that students' tacit beliefs can best be viewed.

In this chapter, I describe in detail the methodology used to elicit students' epistemic criteria. I begin with a description of the study participants and setting. Next, I outline the design of the tasks used in this study and procedures for administering the tasks. Finally, I introduce the analytical frame and coding procedures that I used for the analysis of student responses.

Study Participants

A total of 42 students participated in this study. Twenty-one students were elementary school students (mean age 10.3 yrs SD .31), and they were all entering the fifth grade. Twenty-one students were high school students (mean age 16.9 yrs SD .68). The majority of the high school students were in the eleventh grade (six were in 12th grade). The elementary school cohort included 10 female and 11 male students. The high school cohort includes 11 female and 10 male students.

All students were recruited from a mid-size urban school district in California with a highly diverse[1], multi-ethnic student population with 42% of students qualifying for free or reduced lunch. Elementary school students were recruited from district-run summer school classrooms. Recruiting students from the summer program allowed for a student sample from nine different elementary schools. This diversity mitigated against the influence that particular instruction may have had on students had they all been part of the same classroom at a particular school.

The high school students were all recruited from the only high school in the district. They attended the "green academy" that had a diverse student body generally interested in environmental issues. The high school students were

---

[1] Asian (8%), Latino (22%), African-American (22%), White (33%), Other or Multi-Ethnic (15%)

recruited from several environmental studies classes that were required for all students in the academy. The fact that the classes were mandatory avoided selecting from optional classes with self-selecting students (such as optional advanced courses.)

To recruit students, I briefly introduced myself, the subject of the study (generally framed as "ideas about science") and the time commitment required for participation. All students who were entering 5th grade, who volunteered, and who provided the appropriate permissions, were admitted to the elementary school cohort. All 11th and 12th grade students who volunteered and provided appropriate permissions were admitted to the high school cohort.

Task Design and Procedures

I designed three tasks to elicit students' epistemic criteria about what makes a scientific research question worthwhile. Each task was conducted in a one-on-one semi-structured interview. The interviews were video-recorded and student responses were transcribed for analysis. In this section, I describe each task in detail, the methods for administering the task and the task's affordances for eliciting student thinking. The full interview protocol used to administer the tasks is included in Appendix A.

Funding allocation task

The Funding Allocation Task was designed to elicit students' epistemic criteria as applied to scientific research questions in a funding context. Scientific research requires funding. Funding sources are limited, so scientists must make a convincing case that their research is worthwhile to investigate and thereby merits funding. In the Funding Allocation Task, students were given fake money and presented with 14 scientific research questions as potential questions to fund (Table 1). The questions were designed to reflect a diverse range of issues of varying interest to students. The set of questions included "why," "what," "which," "how" and "yes or no" questions. Some questions targeted trivial topics (e.g., paper towels) and others targeted very serious topics (e.g., earthquakes).

The Funding Allocation Task was designed such that students did not have sufficient funds to allocate funding to all of the scientific research questions. Therefore, students were forced to evaluate the candidate questions and decide which questions to fund, which questions not to fund and how much money to allocate to each of the funded questions.

Table 1: *Candidate Research Questions in Funding Allocation Task*

Q#	Handle	Scientific Research Question
1	TOWEL	Which kind of paper towel is the best at wiping up a spill?
2	SODA	What happens when plants are given soda instead of water?
3	METALS	Why are some metals stronger than other metals?
4	MOON	What causes the moon to change shape?
5	DINO	What color eyes did dinosaurs have?
6	CHEMICALS	What happens when different chemicals are mixed together?
7	EARTHQUAKE	Is it possible to predict when an earthquake will happen?
8	ALIENS	Are there living things on other planets?
9	MAGNET	What happens when you run electricity through a magnet?
10	BEES	Why are bees getting sick?
11	MATH	What is the best way to teach kids math?
12	COMPUTERS	How can we make computers faster?
13	SOCCER	Can boys or girls kick a soccer ball farther?
14	PESTS	How can plants be protected from pests without using chemicals?

At the beginning of the Funding Allocation Task, the interviewer explained that scientists ask questions and conduct investigations to try to answer those questions. The interviewer then explained that scientists need money to conduct their investigations and that scientists apply for funding that is awarded to scientists who are asking *worthwhile* questions. Students were presented with the problem that there is not sufficient funding to fund all of scientists' research questions. The students were then tasked with deciding which questions they thought were worthwhile to receive funding.

Each student was given ten $100 bills and asked to listen carefully to the research questions as the interviewer read each question and showed the question printed on a laminated card. The students were asked to physically move each question to one of two headings that were placed on the table: "Probably Want to Fund" or "Probably Don't Want to Fund." After all 14 questions were read by the interviewer and placed under one of the two headings by the students, the students were then asked if they would like to change their minds about any of the questions. This opportunity was offered to the students in case their funding decisions depended on having seen all of the candidate research questions.

For each question that the students declined to fund, the students were asked to explain their rationale for the decision not to fund. The students were then asked to allocate the $1000 (in $100 bills) across the set of funded questions and to explain, one by one, their reasons for funding each question (Figure 2).

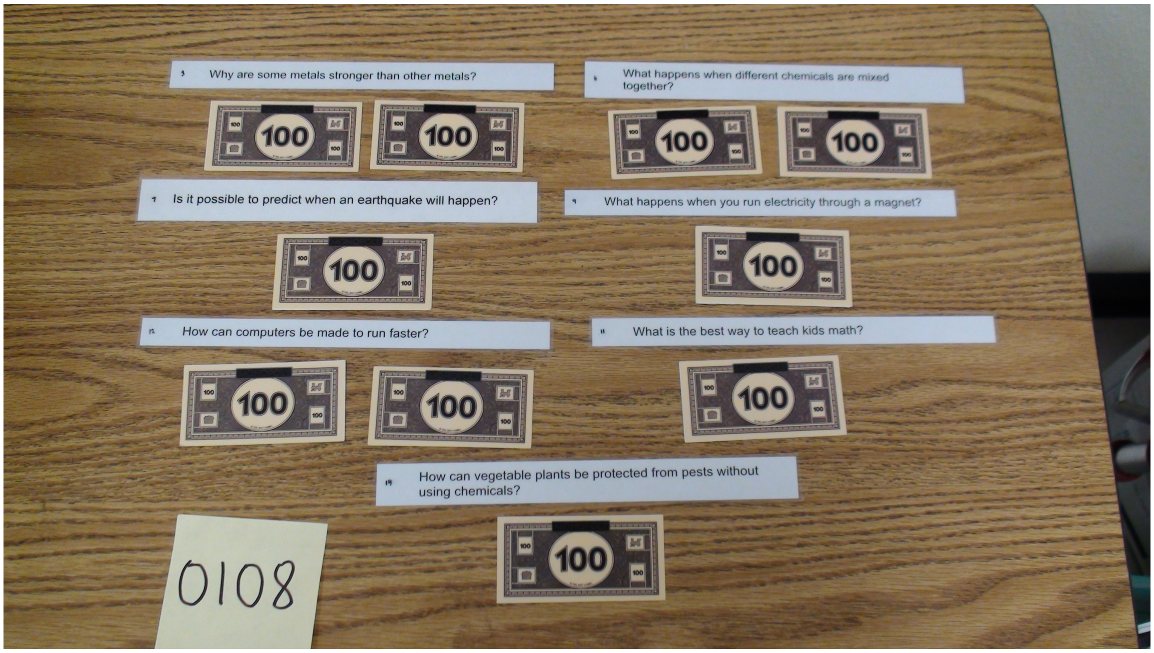


Figure 2: Sample Funding Allocation

After explaining their funding decisions, the students were asked two final questions. First, students were asked which question they would fund if they had to allocate the entire amount ($1000) to a single question. Second, they were asked to explain how they made their funding decisions. This last question prompted students to generalize their thinking about their decision-making process and to articulate the criteria they used to evaluate the scientific research questions. Throughout the task, the interviewer prompted the students to explain their ideas and asked questions to clarify student thinking.

The Funding Allocation Task elicited students' own ideas about what makes a scientific research question worthwhile. The task invited students to make practical judgments about funding questions and to explain their rationales for their decisions. While the task did focus on scientific research questions for investigation by professional scientists, students made their *own* judgments about which questions they considered to be worthwhile rather than what they thought a scientist might think about the questions. Therefore, the task elicited students' epistemic ideas as reflected in their practice (allocating funding) in a context about professional science.

By prompting students to retrospectively reflect on their funding decisions, the task also afforded a comparison between students' "on-line" decisions specific to each question and their retrospective articulation of generalized criteria. These two types of thinking are distinct, and distinguishing between them provides a fuller picture of student thinking. Self-reporting about how students made a decision or solved a problem is a distinct cognitive process, separate from the thinking involved in making the decision or solving the problem (Ericsson & Simon, 1980).

Ladybug task

The Ladybug Task was designed to elicit students' epistemic criteria in relation to the epistemic utility of particular research questions. In contrast to the Funding Allocation Task which included research questions on a wide range of topics (from dinosaurs to earthquakes,) the Ladybug Task invited students to evaluate the merits of a set of questions that all focus on the same topic. The task presented a specific idea that the questions were intended to investigate. Students were expected to decide which questions were worthwhile to investigate, in which order to investigate those questions and a rationale for that order.

The Ladybug Task was administered following the Funding Allocation Task. Once the Funding Allocation Task was completed, the interviewer cleared away all of the funding questions with the exception of one: "How can plants be protected from pests without using chemicals?" Students were asked if they had any ideas about how to protect plants from pests without using chemicals. Next, the interviewer explained that a hypothetical student was really interested in this issue and wanted to conduct a science investigation to help him/her better understand what was happening with the pests in order to find a way to protect the plants without using chemicals.

In order to develop students' content knowledge about ladybugs and aphids, the students were shown a brief video that the hypothetical student captured of his or her garden. The video showed a close-up view of aphids on a plant and ladybugs eating the aphids. The students were asked what they noticed in the video and were told that ladybugs live inside flowers. If they did not already know, the students were also told that aphids are pests that eat broccoli plants.

They were then told that the hypothetical student had an idea that he or she thought might work to protect the plants from the aphids. The students were presented with photo manipulatives of a broccoli plant, a ladybug, a flower and aphids (see Figure 3) and asked what they thought the hypothetical student's idea might have been. The video and photos were designed to expose students to the phenomenology of the garden and to build their content knowledge about ladybugs, broccoli and flowers. At this point in the task, it was important for the students to understand the idea of planting flowers next to the broccoli plants to attract ladybugs that can then eat the aphids. If a student did not clearly explain this idea to the interviewer, the interviewer used the manipulatives to explain the idea to the student.

Figure 3: Ladybug Task Photo Manipulatives

The main task of the student was to determine which of the questions that the hypothetical student generated would make the most sense to investigate if he/she wanted to test his/her idea. The students were presented with a set of five candidate questions on typed, laminated cards (see Table 2) and asked which questions they thought made sense to investigate. The students were invited to select as many questions as they thought made sense to investigate. The interviewer asked the students to explain their reasoning for their decisions. Finally, students were asked to sequence the questions in the order which they thought made the most sense to investigate. Students were asked to explain why they placed the questions in the order that they did. Throughout the task, the interviewer prompted the students to explain their ideas and asked questions to clarify student thinking.

Table 2: Candidate Research Questions in Ladybug Task

Question #	Research Question
Q1	How many aphids can a ladybug eat in one day?
Q2	Does planting flowers next to broccoli cause fewer aphids?
Q3	How many spots do ladybugs have on their backs?
Q4	Which flowers do ladybugs like best?
Q5	Which part of broccoli plants do aphids eat first?

The Ladybug Task invited students to evaluate scientific research questions on the same topic in relation to a particular idea of interest. The task scaffolded students' content knowledge in the topic so students would have some knowledge upon which to ground their reasoning about particular research questions. The task also invited students to reason about contingency between multiple questions and how those questions could be coordinated with each other. Coordinating multiple research questions is both typical of authentic science research programs and highlights how scientific knowledge is often the product of many different scientific investigations.

Science fair task

The Science Fair Task was designed to elicit students' epistemic criteria about scientific research questions in the context of school science and a science fair in particular. The science fair is a common school setting in which students share their science projects with their classmates, parents and school community. It is common for particularly strong science fair projects to receive awards. The criteria that judges use at science fairs to award prizes is important to consider because they reflect the aspects of science valued by school science. In the Science Fair Task, students were asked which criteria they thought the science fair judges should use when awarding the prize for "Best Scientific Research Question"(Figure 4). The students' responses to this task provide a view of what they think makes a worthwhile scientific research question in a school setting. Thus the task affords an analysis of students' epistemic criteria to determine the extent to which students' criteria depend on the context (professional vs. school).

Figure 4: Science Fair Task Materials

At the beginning of the Science Fair Task, the interviewer told the students that a school was having a science fair. The students were asked if they knew what a science fair is and if they had ever been to a science fair. The interviewer explained that a science fair is an event where students present the scientific research projects that they have worked on. The interviewer also explained that, prior to the science fair, each student in the grade was required to develop a scientific research question that they wanted to investigate and then they each conducted their investigation. Next, the interviewer explained that science fair judges gave out a special award for the "Best Scientific Research Question." The judges were given instructions that explained what makes a scientific research question worthwhile and how they should make their decision. The students were asked what instructions they would have given to the science fair judges. This task invited students to share the epistemic criteria for scientific research questions that they believed should be applied in school contexts.

Analytical Frame

The analyses in this dissertation focus on the responses that students provided as they engaged in the three interview tasks. Figure 5 illustrates how each task contributed data to the different analyses in the subsequent chapters of the dissertation.

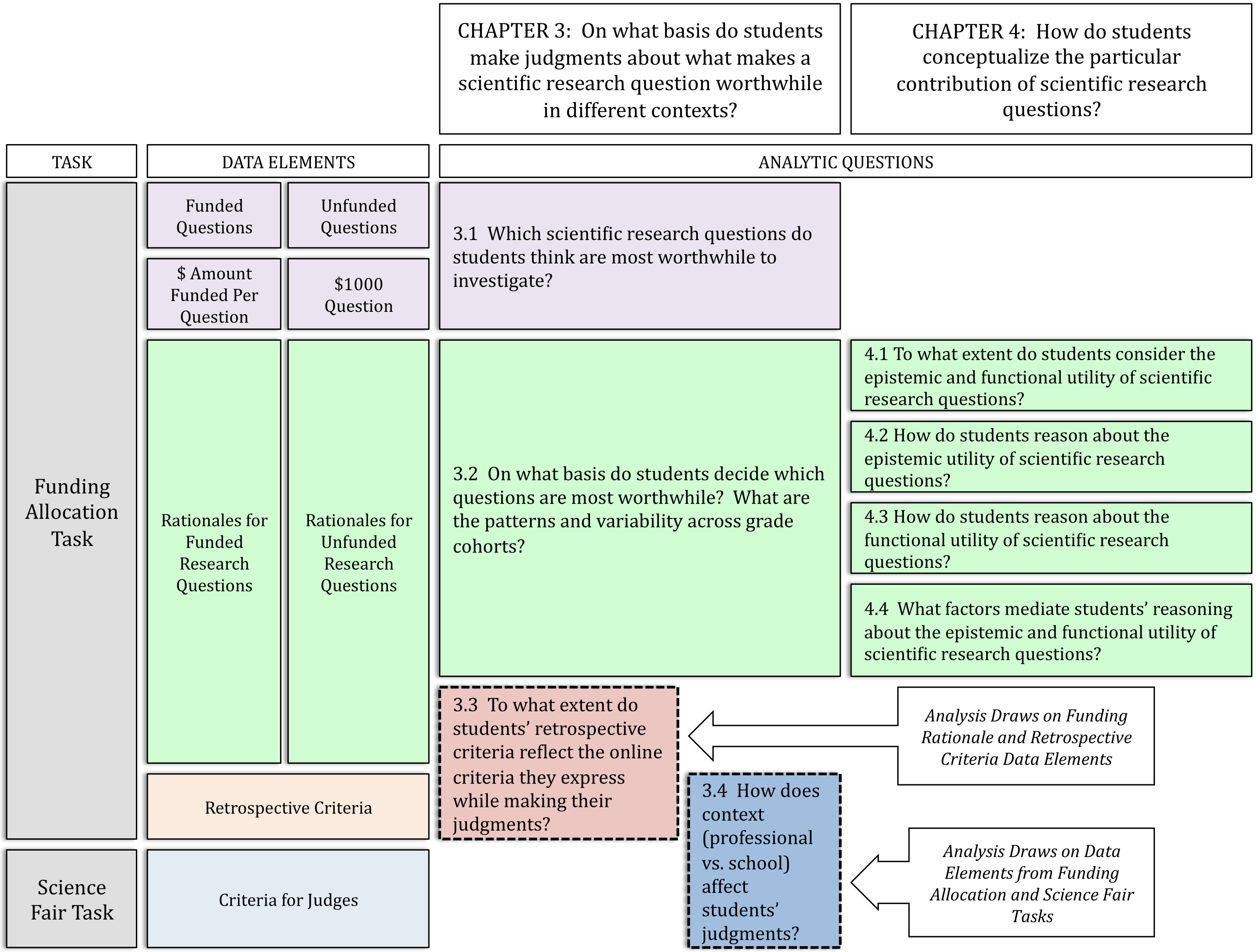
In Chapter 3, I present analysis to address the first research question of the dissertation: On what basis do students make judgments about what makes a scientific research question worthwhile in different contexts? To address this question, I focus on four subordinate questions. First, I analyze which scientific research questions students find most worthwhile to investigate. This analysis draws on several student responses from the Funding Allocation Task including the questions students fund, the questions students don't fund, the amount of funding each question receives and the question students select as the $1000 question.

Second, I analyze the basis on which students make their decisions about which questions are worthwhile to investigate. This analysis is based on the rationales that students provide for their decisions about which questions to fund and which questions to not fund.

In the third analysis, I examine the extent to which students' retrospective criteria reflect their on-line judgments about specific questions. This analysis compares the retrospective criteria and students' judgments about their individual funding decisions. The final analysis examines how students' epistemic criteria differ by context. This analysis compares students' retrospective criteria from the Funding Allocation Task with the criteria students recommend that the judges use in the Science Fair Task.

In Chapter 4, I present analyses to address the second research question of the dissertation: How do students conceptualize the particular contribution of scientific research questions? To address this question, I focus on four subordinate questions. First, I analyze the extent to which students consider the epistemic and functional utilities of the research questions in the Funding Allocation Task based on the rationales the students provide for their funding decisions. Next, I analyze specific patterns in students' rationales about how the students think about the epistemic utility and functional utility of research questions. Finally, I examine particular factors that influence how students conceptualize the utility of scientific research questions. All of the analyses in Chapter 4 draw on students' funding rationales in the Funding Allocation Task.

In Chapter 5, I present analyses to address the third research question of the dissertation: How do students evaluate and coordinate a set of scientific research questions to investigate an idea? To address this question, I explore four analytic questions using data derived from the Ladybug Task. The first analysis examines which questions students thought made most sense to investigate. The second analysis focuses on the rationales that students provided for why they thought certain questions made sense to investigate. The third analysis examines the variation in students' question sequences and the final analysis looks at the rationales that students provide for their question sequences. Together, these analyses examine the extent to which students think about the epistemic utility of scientific research questions and the epistemic contingency involved in considering different potential question sequences.



CHAPTER 3: On what basis do students make judgments about what makes a scientific research question worthwhile in different contexts?

CHAPTER 4: How do students conceptualize the particular contribution of scientific research questions?

3.1 Which scientific research questions do students think are most worthwhile to investigate?

3.2 On what basis do students decide which questions are most worthwhile? What are the patterns and variability across grade cohorts?

3.3 To what extent do students' retrospective criteria reflect the online criteria they express while making their judgments?

3.4 How does context (professional vs. school) affect students' judgments?

4.1 To what extent do students consider the epistemic and functional utility of scientific research questions?

4.2 How do students reason about the epistemic utility of scientific research questions?

4.3 How do students reason about the functional utility of scientific research questions?

4.4 What factors mediate students' reasoning about the epistemic and functional utility of scientific research questions?

Figure 5: Analytical Frame

Chapter 5: How do students evaluate and coordinate a set of scientific research questions to investigate an idea?

TASK	DATA ELEMENTS		ANALYTIC QUESTIONS
Ladybug Task	Questions that Make Sense to Investigate	Questions that Don't Make Sense to Investigate	5.1 Which questions do students think make sense to investigate?
	Rationales for Questions that Make Sense to Investigate	Rationales for Questions that Don't Make Sense to Investigate	5.2 What utility do students attribute to different scientific research questions?
	Question Sequence		5.3 How do students sequence research questions to test an idea?
	Rationale for Question Sequence		5.4 On what basis do students rationalize the coordination of research questions?

Figure 5: Analytical Frame (Continued)

Procedures for Data Analysis

Student responses to the Funding Allocation, Ladybug and Science Fair Tasks were the data sources for the analyses in the following chapters of the dissertation. In this section, I provide an overview of the process and methods for analyzing the student data. I describe how the coding schema was developed, the coding process and the inter-rater coding process. Further descriptions of the analyses are presented in each analytic chapter (Chapters 3, 4 & 5).

The interview produced student response data that could be analyzed in a straightforward manner as well as response data that required more involved interpretation. Straightforward data in the Funding Allocation Task included (1) the specific questions that were funded or not funded, (2) the amount of money given to each question, and (3) the question that the student selected to receive the full funding of $1000. For the Ladybug Task, the straightforward data included the specific decision of whether a question made sense to investigate or not (either "made sense" or "did not make sense"). The different orders of question sequences in the Ladybug Task were organized according to permutation.

Data such as students' rationales and criteria required more interpretation and thus necessitated the creation of specific schemas for coding and analysis purposes. Three distinct coding schemas were developed to interpret students' rationales and criteria (see Table 3). The coding schemas were developed to reflect patterns in students' explanations. The schemas were designed to function as a set of distinct codes that reflected the full range of student responses.

All three coding schemas were derived emically through an iterative review of student responses. To develop each coding schema, I conducted a preliminary review of the range of student responses. Student response types that were found repeatedly were assigned a particular code in the emerging coding schema. As the coding schemas emerged, I applied them to the data set and continuously adjusted them until they adequately captured student thinking. The final coding schemas were then used to score the complete data set. Individual codes were also grouped together into larger clusters of codes to facilitate analysis. The coding schemas themselves are presented in the relevant analytical chapter (Chapters 3, 4 & 5).

Table 3: Schemas for Coding Different Task Components

Task	Student Responses	Coding Schema
Funding Allocation Task	Student rationales for funding decisions	Rationale & Criteria Coding Schema (Chapters 3 & 4)
	Student retrospective criteria	
Science Fair Task	Student criteria for judges	
Ladybug Task	Student rationales for questions that made sense or did not make sense to investigate	Epistemic Utility Coding Schema (Chapter 5)
	Student rationale for the question sequence	Sequence Rationale Coding Schema (Chapter 5)

Coding process

All student interviews were video-recorded and transcribed. Each student transcript was parsed into units of analysis for coding. The unit of analysis consisted of a rationale related to a particular question or a specific criteria listed by the students. Students often provided more than one rationale for each evaluation of a particular scientific research question. Each rationale was assigned an individual code so students' explanations for why they funded a particular question could receive more than one code. For prompts that asked students to list criteria (retrospective funding criteria and science fair criteria) each criterion was assigned its own code.

A research assistant served as a second coder. After being trained to use the coding schema, the assistant coded a randomly selected sample (20% of the students) using the Rationale & Criteria Coding Schema. Two coders coded the "on-line" funding rationales, retrospective criteria and science fair criteria for the coding sample. Matching codes between the two coders was particularly challenging due to the numerous possible coding combinations. For example, a single student's explanation for a particular funding decision could include a variable number of rationales and each rationale could be assigned one of 30+ individual codes.

Therefore, coding "agreement" took two different forms. Either the two coders picked identical codes or the two coders picked different codes within the same cluster (see Chapter 3 for coding clusters). There were also two forms of coding "disagreement." Either the two coders selected codes from different clusters or one coder assigned a code and the other coder did not. Disagreements were resolved through discussion between the primary and secondary coders.

The inter-rater reliability was calculated as the percentage of agreements out of the total (sum of all agreements and disagreements). Table 4 shows the percentage of agreements for each component of the interview. While modest, the percent agreements were reasonable given the remarkable variability and flexibility in the quantity of students' rationales and criteria and the wide range of possible codes to assign.

Table 4: Inter-Rater Percent Agreement

	On-line Funding Rationales	Retrospective Criteria	Science Fair Criteria
Inter-Rater Percent Agreement	81%	68%	73%

In the subsequent chapters, I analyze student responses to the three tasks using the analytical approach outlined above. In each analytical chapter, I provide more details about the specific coding procedures that are particular to each analysis. The next three chapters describe the analyses and empirical findings derived from student responses to the interview tasks outlined in this chapter. Chapter 3 focuses on analyses drawing on the Funding Allocation and Science Fair Tasks. Chapter 4 includes findings from a finer grain analysis of student responses to the Funding Allocation Task. Chapter 5 focuses on analysis of student responses to the Ladybug Task.

Chapter 3: On What Basis Do Students Form Judgments About What Makes a Scientific Research Question Worthwhile?

This chapter examines how students evaluate what makes a scientific research question worthwhile to investigate. The analysis focuses primarily on the criteria that students use when they evaluate the "worthwhileness" of a scientific research question. The practice of determining "what counts" as a worthwhile research question is an epistemic practice. Therefore, students' responses reflect their practical epistemology (Sandoval, 2005) as applied to scientific research questions. The assessment tasks analyzed in this chapter are specifically designed to elicit students' epistemic criteria in the context of practice.

In this chapter, I analyze student responses to two different tasks that target students' epistemic reasoning. The "Funding Allocation Task" is a novel task designed to elicit students' judgments about which scientific research questions are worthwhile for professional scientists to investigate. In the task, students are asked to evaluate a set of 14 scientific research questions posed by professional scientists. Students are given $1000 in pretend money (ten $100 bills) to allocate to the questions they think are worthwhile for scientists to investigate and then to explain their rationales for their funding decisions. At the end of the task, students are asked to explain retrospectively which criteria they used to make their funding decisions.

The "Science Fair Task" is designed to elicit students' criteria for what counts as a worthwhile scientific research question for a school science fair project. Students are asked what criteria they think judges at a science fair should consider when deciding how to award the prize for "Best Scientific Research Question." Students respond by listing the criteria they think are important to consider. By asking students to consider criteria for a worthwhile research question in a science fair context, we can see if students apply different criteria to research questions investigated by professional scientists compared to those investigated by science students.

The analysis in this chapter is organized around the following questions:

1. Which scientific research questions do students think are most worthwhile to investigate?
2. On what basis do students decide which questions are most worthwhile? What are the patterns and variability across grade cohorts?
3. To what extent do students' retrospective criteria reflect the on-line criteria they express while making their judgments?
4. How does context (professional vs. school) affect these judgments?

Which Scientific Research Questions Do Students Think Are Most Worthwhile to Investigate?

In the "Funding Allocation Task" students were presented with 14 candidate scientific research questions and $1000 worth of funding to allocate. Table 5 lists the 14 candidate research questions. Students were asked to select the research questions they found most worthwhile and to allocate funds to those questions. Students were then asked to explain their funding decision for each of the 14 research questions.

Table 5: Candidate Research Questions in Funding Allocation Task

Q#	Handle	Scientific Research Question
1	TOWEL	Which kind of paper towel is the best at wiping up a spill?
2	SODA	What happens when plants are given soda instead of water?
3	METALS	Why are some metals stronger than other metals?
4	MOON	What causes the moon to change shape?
5	DINO	What color eyes did dinosaurs have?
6	CHEMICALS	What happens when different chemicals are mixed together?
7	EARTHQUAKE	It is possible to predict when an earthquake will happen?
8	ALIENS	Are there living things on other planets?
9	MAGNET	What happens when you run electricity through a magnet?
10	BEES	Why are bees getting sick?
11	MATH	What is the best way to teach kids math?
12	COMPUTERS	How can we make computers faster?
13	SOCCER	Can boys or girls kick a soccer ball farther?
14	PESTS	How can plants be protected from pests without using chemicals?

In this section, I analyze the decisions that students made about whether to fund or not to fund the scientific research questions. While subsequent analyses focus on criteria students used to make their decisions, I begin with an analysis of the questions themselves. To determine which questions the students considered to be most worthwhile, I analyze and triangulate three indicators: funding rate, funding amount and the one question they would fund with their entire $1000 ("$1000 question"). Using each of these indicators, I identify the research questions deemed most worthwhile and least worthwhile according to students, including differences between grade cohorts. Following an analysis of the indicators, I summarize the results across all three indicators.

Funding rate

The funding rate is the most basic indicator of the questions that students considered most and least worthwhile. The funding rate for each question represents the percentage of students who chose to fund that question. Table 6 shows the funding rate for all students, elementary school (ES) students and high school (HS) students.

Table 6: Funding Rate by Question and Grade Cohort

Question	% Funding Rate (ES)	% Funding Rate (HS)	% Funding Rate (All)
Q1-TOWEL	24%	0%	12%
Q2-SODA	33%	10%	21%
Q3-METALS	38%	57%	48%
Q4-MOON	67%	33%	50%
Q5-DINO	52%	19%	36%
Q6-CHEMICALS	62%	62%	62%
Q7-EARTHQUAKE	76%	71%	74%
Q8-ALIENS	62%	71%	67%
Q9-MAGNET	38%	38%	38%
Q10-BEES	43%	86%	64%
Q11-MATH	57%	71%	64%
Q12-COMPUTERS	57%	48%	52%
Q13-SOCCER	33%	10%	21%
Q14-PESTS	76%	95%	86%

It is worth noting the wide range of funding rates across the different research questions (overall funding rates range from 12% to 86%). Furthermore, the funding rates are distributed quite evenly across that range. This indicates that the research questions were sufficiently controversial. In other words, the variation in funding rates means that, for each question, there was a reasonable portion of students who funded and did not fund each question. The absence of widespread consensus means that student judgments were not uniform, nor are the candidate questions obviously fundable or non-fundable. Therefore, the conceptualizations of a worthwhile question are not straightforward and warrant further analysis to account for the observed variation.

Funding rates in Table 6 indicate that Q14-PESTS and Q7-EARTHQUAKES were funded by the highest percentage of students and Q13-SOCCER, Q2-SODA and Q1-TOWEL had the lowest funding rates. There were also three research questions that had particularly large gaps in funding rates between grade cohorts (see Figure 6). The HS funding rate (86%) for Q10-BEES was twice as high as the ES funding rate (43%). Q4-MOON and Q5-DINO had much higher funding rates for the ES cohort compared to the HS cohort.

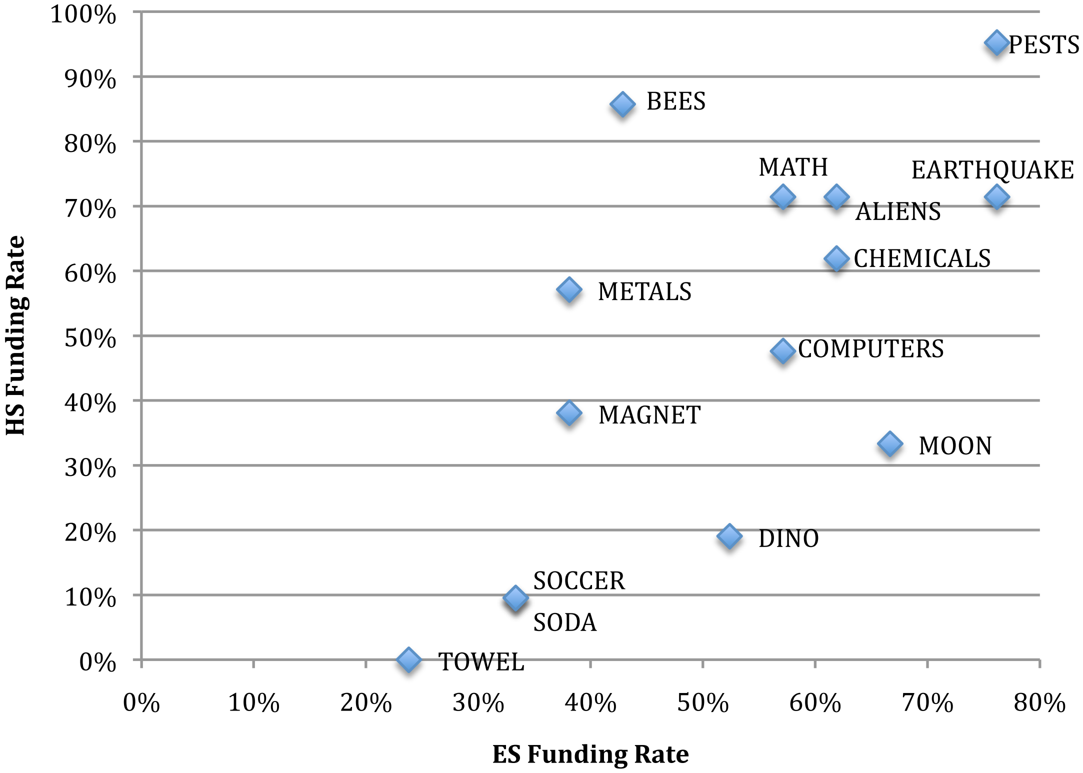


Figure 6: Funding Rates by Grade Cohort

Funding amount

Another indicator of which research questions students thought were most worthwhile is the amount of money that students allocated to each research question. As a complement to the funding rate, the funding amount takes into account the "strength" of the endorsement. A higher funding amount is interpreted as a stronger endorsement of the research question. Figure 7 shows the *per student* funding allocations for each research question by grade cohort. Among all students, the highest per student funding amounts are for the Q7-EARTHQUAKE and Q8-ALIENS questions. Q14-PESTS received almost as much funding from students. These funding amounts corroborate the funding rate findings.

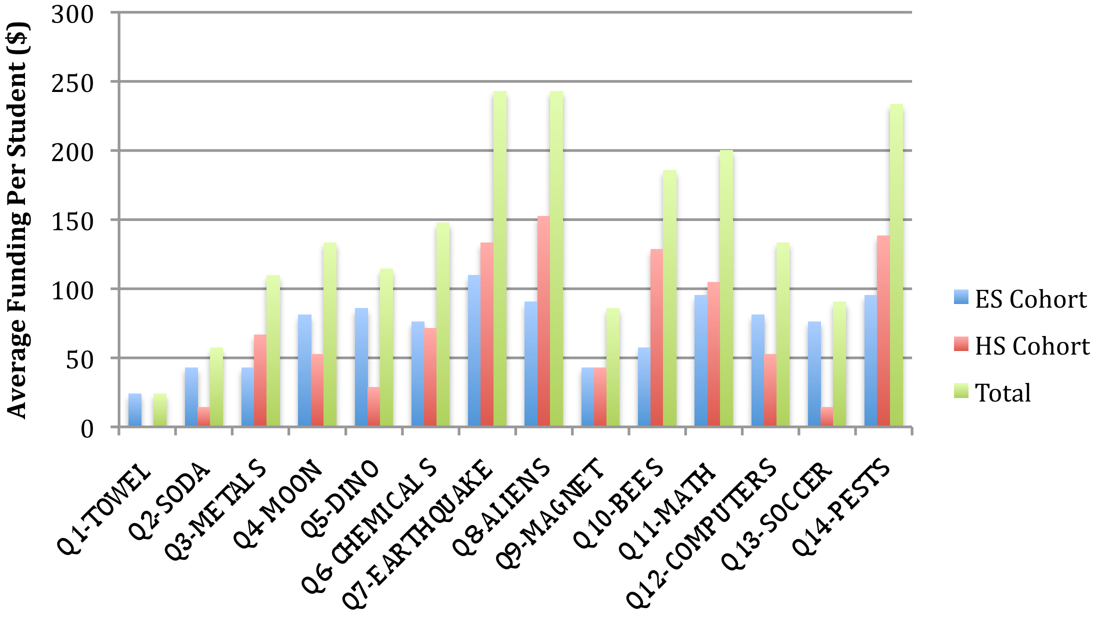


Figure 7: Per Student Funding Amounts by Research Question and Grade Cohort

Several research questions had large funding gaps between the ES and HS grade cohorts. The largest funding gaps were for Q10-BEES, Q5-DINO and Q8-ALIENS. While the funding amounts for Q10-BEES and Q5-DINO are expected given the differences in funding rates, the funding gap for Q8-ALIENS indicates that the HS students not only funded this question at a higher rate than the ES students, but they also funded at a disproportionally higher amount.

$1000 question

The last indicator of which questions students considered most worthwhile is the students' $1000 allocation. At the end of the Funding Allocation Task, after students had allocated their funds and explained their decisions, each student was asked which question they would fund if they had to give all $1000 to a single question. In other words, which question is the *most* worthwhile?

Figure 8 shows a frequency count of students and the question they chose to award $1000. Nine of the 14 research questions were selected by at least one student. At least one ES student selected each of the nine questions, whereas the HS students only selected six different questions. This is another indication that the set of research questions was diverse and appealed to different students in different ways. Q7-EARTHQUAKE and Q8-ALIENS were the two most commonly selected questions overall and the most common in each grade cohort. Q5-DINO, Q12-COMPUTERS and Q13-SOCCER were only selected by students in the ES cohort.

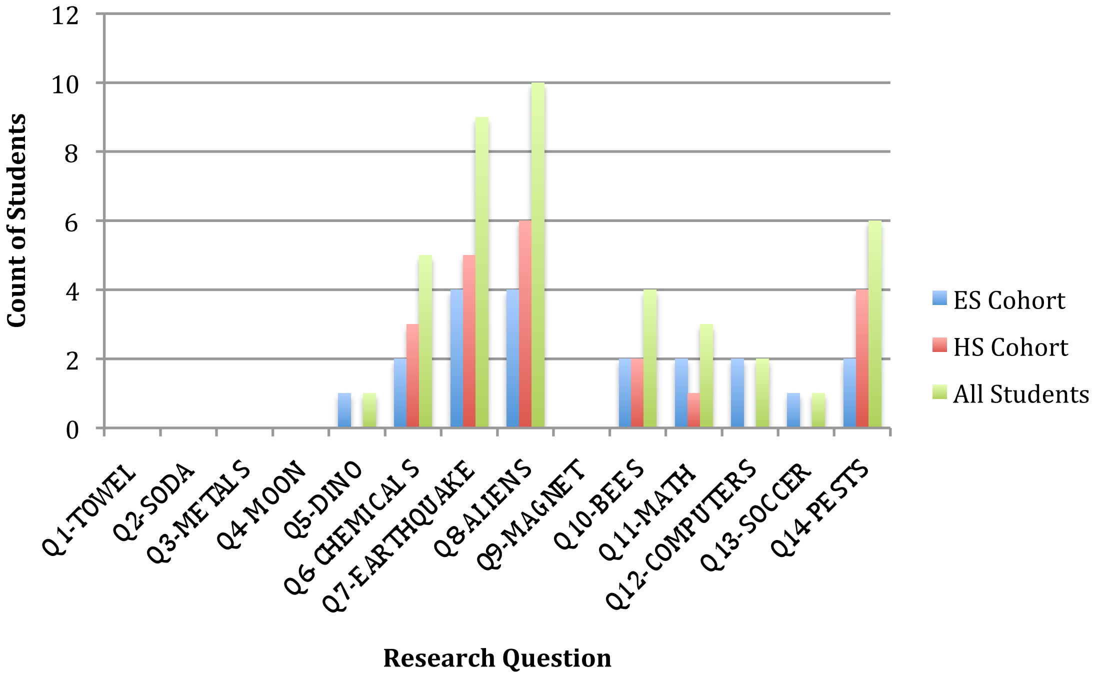


Figure 8: Frequency of $1000 Funded Questions by Grade Cohort

Summary of indicator findings

Taken together, the funding rate, funding amount and $1000 question selection all serve as indicators of which questions students thought were most worthwhile. Table 7 summarizes the findings from these indicators. Q7-EARTHQUAKES was rated highly across all three indicators. Q8-ALIENS and Q14-PESTS rated highly in two of the three indicators. These findings indicate that students considered Q7-EARTHQUAKES, Q8-ALIENS and Q14-PESTS as the most worthwhile research questions for scientists to investigate. Similarly, Q1-TOWEL and Q2-SODA were rated low across all three indicators. Students considered these two research questions to be the least worthwhile. Finally, there were large grade cohort gaps in the ratings for Q4-MOON, Q5-DINO and Q10-BEES. Grade cohort differences for the $1000 question were minor (maximum difference of two students).

While these analyses are useful for identifying *which* scientific research questions students considered most worthwhile, further analysis is needed to understand *why* students made these judgments. It is possible that students viewed Q13-SOCCER, Q2-SODA and Q1-TOWEL as questions very familiar or straightforward and therefore not worth the effort of professional scientists.

Table 7: Summary of Worthwhile Question Indicators

Indicator	Highest Rate, Highest Amount or Most Frequent	Lowest Rate, Lowest Amount or Least Frequent	Gap Between ES & HS Cohorts
Funding Rate	Q7-EARTHQUAKES Q14-PESTS	Q1-TOWEL Q2-SODA Q13-SOCCER	Q4-MOON *(x2 higher for ES)* Q5-DINO *(x2.7 higher for ES)* Q10-BEES *(x2 higher for HS)*
Funding Amount	Q7-EARTHQUAKE Q8-ALIENS Q14-PESTS	Q1-TOWEL Q2-SODA	Q5-DINO *(x3 higher for ES)* Q8-ALIENS *(1.7x higher for HS)* Q10-BEES *(2.25x higher for HS)*
$1000 Question	Q7-EARTHQUAKE Q8-ALIENS	Q1-TOWEL Q2-SODA Q3-METALS Q4-MOON Q9-MAGNET	Q8-ALIENS *(2 more HS students)* Q12-COMPUTERS *(2 more ES students)* Q14-PESTS *(2 more HS students)*

One possible explanation for the grade cohort gaps is differences in content knowledge. High school students may be more aware of the role bees play in our ecosystem and thus are more likely to recognize bee illness as a cause for concern and worthy of investigation. Differences in moon content knowledge between the two grade cohorts might explain the different funding rates for Q4-MOON. If HS students know why the moon changes shape, it makes sense that they would be less likely to consider that question worthwhile. Q5-DINO may have appealed to younger students who may be more captivated by dinosaurs compared to older students. In the next section, I analyze the types of rationales that students gave to explain their funding decisions. The analysis will help explain the funding patterns encountered thus far.

On What Basis Do Students Decide Which Questions Are Most Worthwhile? What Are the Patterns and Variability Across Grade Cohorts?

I began with an analysis of which scientific research questions students considered most worthwhile. I now turn to an analysis of the rationales students gave for their funding decisions. By focusing on student rationales, I can determine the criteria that students applied when they were deciding what makes a scientific research question worthwhile to investigate. In this section, I begin by presenting a coding scheme to interpret student rationales. I then analyze patterns in the rationales that students gave for their judgments of the 14 scientific research questions. I attend to patterns within and across grade cohorts.

Coding schema and procedure

To interpret student rationales, I developed an *emic* coding schema using an iterative process. In a preliminary review of students' rationales, I identified instances of reoccurring rationales and I assigned codes to each rationale type that occurred more than once. As the schema emerged, I applied the schema to the data, making adjustments as necessary so that the schema adequately captured the rationales stated by the students. The individual codes are organized into clusters as shown in Table 9.

Table 8: Coding Clusters for Student Rationales in Funding Allocation Task

Coding Cluster	Cluster Description
Knowledge Status	Student rationales based on whether the research question is known or unknown
Difficulty	Student rationales based on the perceived difficulty of the research question to investigate
Interest	Student rationales based on interest in the research question
Contribution	Student rationales based on the expected contribution of the research question
Scientific Attributes	Student rationales based on the extent to which the research questions have attributes of scientific questions
Non-Response	Assigned to cases in which the student answers the research question itself instead of explaining basis for decision or otherwise uncodeable

The first coding cluster, "Knowledge Status," includes codes that pertain to student rationales that attribute their funding decision to whether the question is already known (Table 9). The set of codes in this cluster reflects students' rationales indicating that the question is "already known" or "unknown" by the student, scientist or people in general, or if the knowledge can be found in books or on the internet.

The "Difficulty Cluster" pertains to student rationales that consider how hard the question would be to investigate (Table 9). Some questions are very difficult to investigate (e.g., DIF, NPOS), whereas others are fairly simple and straightforward (e.g., EA, DIY).

The "Interest Cluster" pertains to student rationales that consider the extent to which the research question is interesting (Table 9). This cluster includes codes for whether the question is interesting to the student (INTS or NINTS) and interesting to people in general (INTP). Questions that evoke an emotional response (EMO) are also included in this cluster.

Table 9: Individual Codes for Student Rationales by Cluster

Knowledge Status Cluster

Code	Code Name	Description	Example(s)
UKS	Unknown (by student)	Student explains funding decision based on fact that he/she does not know the answer to the research question.	"I never learned about it"
AKS	Already known (by student)	Student explains funding decision based on fact that he/she already knows the answer to the research question.	"already studied"
UKSC	Unknown (by scientists)	Student explains funding decision based on fact that scientist(s) do not know the answer to the research question.	"scientists don't know"
AKSC	Already known (by scientist)	Student explains funding decision based on fact that scientist(s) already know the answer to the research question.	"scientists know"
UKP	Unknown (by people in general)	Student explains funding decision based on fact that people do not know the answer to the research question.	"nobody knows"
AKP	Already known (by people in general)	Student explains funding decision based on fact that people already know the answer to the research question.	"you could just ask a teacher"
AKIB	Already known (on internet or in books)	Student explains funding decision based on fact that answer to the research question can already be found on the internet or in books.	"they could go to like the internet"

Difficulty Cluster

Code	Code Name	Description	Example(s)
EA	Easy	Student explains funding decision based on how easily the research question can be investigated.	"simple to answer" "kind of easy" "most everybody could figure it out"
DIF	Difficult	Student explains funding decision based on how difficult it would be to investigate the research question.	"most difficult"
POS	Possible to Investigate	Student explains funding decision based on the extent to which it is possible to investigate the research question.	"because you could do it"
NPOS	Not Possible to Investigate	Student explains funding decision based on the extent to which it is possible to investigate the research question.	"I don't think there is any way to make it"
DIY	Do It Yourself	Student explains funding decision based on the fact that he/she could investigate the research question him/herself.	"you could just do this at home"
EFF	Effort	Student explains funding decision based on the effort required to investigate the research question.	"took a lot of time" "worked hard on it" "put effort" "tried their best"

Interest Cluster

Code	Code Name	Description	Example(s)
INTS	Interesting to Student	Student explains funding decision based on the extent to which the research question is interesting to the student.	“cool to know” “interesting that we can learn about” “I’m just curious”
NINTS	Not Interesting to Student	Student explains funding decision based on the extent to which the research question is interesting to the student.	“I don’t care” “not interesting”
INTP	Interesting to People	Student explains funding decision based on the extent to which the research question is interesting to people in general.	“people want to know”
NINTP	Not Interesting to People	Student explains funding decision based on the extent to which the research question is interesting to people in general.	“people don’t care about that”
EMO	Emotional	Student explains funding decision based on the extent to which the research question evokes an emotional response.	“proud to do it” “inspires” “enthusiastic” “really psyched”

Contribution Cluster

Code	Code Name	Description	Example(s)
PRS	Personal Benefit	Student explains funding decision based on the anticipated potential benefit to the student resulting from the investigation. Assign if student describes a personal scenario that would be aided by the investigation.	“my computer is slow”
SOC	Social Benefit	Student explains funding decision based on the anticipated potential benefit to society (or people generally) resulting from the investigation. Benefit to people can be implied (i.e. something positive, desirable or good will happen)	“people can be safe”
ENV	Environmental Benefit	Student explains funding decision based on the anticipated potential benefit to the environment, planet or natural world resulting from the investigation.	“help the planet” “help plants and animals” “ecosystem”
NEW	Something New	Student explains funding decision based on something new that may result from the investigation.	“it could make something new”
IMP	Important	Student explains funding decision based on the extent to which the research question is important.	important for us” “big deal” “serious”
DM	Doesn’t Matter	Student explains funding decision based on assertion that the research question is trivial, pointless or otherwise doesn’t matter.	“doesn’t matter” “pointless”
DET	Detrimental Outcome	Student explains funding decision based on assertion that investigating the research question will potentially result in a negative or detrimental outcome.	“could lead to stereotyping”

FP	Faulty Premise	Student explains funding decision based on an assertion that the research question is based on a faulty premise. The student challenges the assumptions of the question itself.	“gender doesn’t matter” “all metals are the same”
SUP	Superior Question	Student explains funding decision based on suggestion of a different question that he/she deems to be more worthwhile.	“maybe if it was something like what resources did they use, did they have a big ecological footprint”
REL	Relevant	Student explains funding decision based on the extent to which the research question is relevant or applicable to our daily lives.	“can happen any time” “based on everyday life”
NREL	Not Relevant	Student explains funding decision based on the extent to which the research question is relevant or applicable to our daily lives.	“not relevant to our lives”
GTK	Good to Know	Student explains funding decision based on the extent to which the research question is good for the student or others to know.	“we should have basic understanding”
TTO	To Tell Others	Student explains funding decision because the student wants to tell others the outcome of the investigation.	“so you could tell people”

Scientific Attributes Cluster

Code	**Code Name**	**Description**	**Example(s)**
SCI	Scientific	Student explains funding decision based on the extent to which the research question is scientific, about a scientific topic or something that scientists would investigate.	“has to do with science” “what a scientist would do” “a good topic”
TEST	Testable	Student explains funding decision based on the extent to which the research question is testable.	“you can prove it”
EVI	Evidence Generating	Student explains funding decision based on the extent to which the research question can generate evidence, data or “proof.”	
BIA	Biased	Student explains funding decision based on the extent to which the research question contains subjective bias.	“is unbiased”

The “Contribution Cluster” pertains to student rationales that consider the contribution that the research question would make (Table 9). There are many different ways that students conceptualized the nature of the question’s contribution. Some codes reflect rationales that refer to contributions in a very general sense (i.e. REL, IMP, GTK, TTO). Other codes reflect more specific contributions that the research question would make to the student as an individual (PRS), the society as a whole (SOC) or the environment (ENV).

In many cases, students decided not to fund a research question because the question “didn’t matter” (DM) or would have some kind of detrimental impact

(DET). The "faulty premise" (FP) code describes cases in which students objected to the premise of the research question itself. For example, some students explained their decision not to fund Q13-SOCCER because they did not think that gender was a relevant factor to consider in soccer ability. In other words, they disagreed with the premise of the question. While the analysis in this chapter focuses on the "Contribution Cluster" as a whole, Chapter 4 examines patterns in the specific "Contribution" codes in more detail.

The "Scientific Attributes Cluster" pertains to student rationales that are based on some type of scientific attribute of the research question (Table 9). Some students based their funding decisions on whether or not the research question is "scientific" (SCI), "testable" (TEST), "generates evidence" (EVI) or is "biased" (BIA). These students are evaluating scientific research questions using specific attributes that they associate with science research questions.

Scoring procedure

The Funding Allocation Task was administered in a one-on-one interview with each student participant. The interview was videotaped and transcribed. The transcription of each student's rationale for each funding decision for each research question was assigned individual codes using the coding schema. For most clusters, a maximum of one code was assigned per rationale and no code was assigned for that cluster if there was no applicable code. Coding for the "Contribution Cluster" mainly followed the same guidelines, except that DET, FP and SUP were eligible to be second codes if they were applicable. If two "Contribution" codes were relevant, the more specific code was used instead of the more general one. For example, if a student said that the research question was important (IMP) and that it helped people (SOC), then only the SOC code was assigned because it is a more specific characterization of the contribution. A total of 712 individual codes were assigned to students' funding decision rationales (n=42 students).

Cluster counts and weights

In this section I describe the analytic procedure that was used to aggregate the codes assigned to students' funding decision rationales. For each student, I first tallied the number of assigned codes across all of the student's rationales for all 14 funding decisions. I then totaled the number of total codes per cluster. Next, I calculated the relative "weight" of that cluster by dividing the number of codes in that cluster by the total number of codes for that student. For example, if a student had a total of 25 codes across all of the funding decisions and five of those codes were in the "Interest Cluster," then that student had an "Interest Cluster" weight of .20 (5 out of 25). This analytic step functions to characterize the proportion of the student's overall criteria that pertains to the particular cluster (in this case "Interest"). The sum of all of the cluster weights equals 1.00 for each student.

To develop aggregate cluster weights across groups of students, I averaged the individual student cluster weights. Table 10 shows the aggregated cluster weights for all of the students in the study. The largest weight was .54 for the

"Contribution Cluster" which means that 54% of all assigned codes were codes from the "Contribution Cluster." Thus, more than half of the criteria that students applied to their funding decisions related to the expected contribution the research question would make. Students consider "Knowledge Status" and "Interest" to approximately the same degree and "Difficulty" about half as much.

Table 10: Aggregate Cluster Weights (All Students)

	Knowledge Status	Difficulty	Interest	Contribution	Science Attributes	Non-Response
Aggregate Cluster Weights (All Students)	0.18	0.07	0.15	0.54	0.02	0.04

While this view of aggregated cluster weights indicates a range of applied criteria, there are limitations of this analysis. It is possible that the relative cluster weights are, in part, a function of the specific set of research questions presented to students. If the task used a different set of research questions instead, it is possible that the cluster weights would be different. Therefore, I now turn to an analysis of patterns and variability across grade cohorts.

Patterns and variability across and within grade cohorts

In this section, I explore the extent to which younger and older students apply different criteria when evaluating the "worthwhileness" of a scientific research question. To conduct this analysis, I compare the cluster weights between two grade cohorts: elementary and high school students. Both grade cohorts were given the same set of research questions. Therefore, the questions are held constant allowing for meaningful comparisons in cluster weights between grade cohorts.

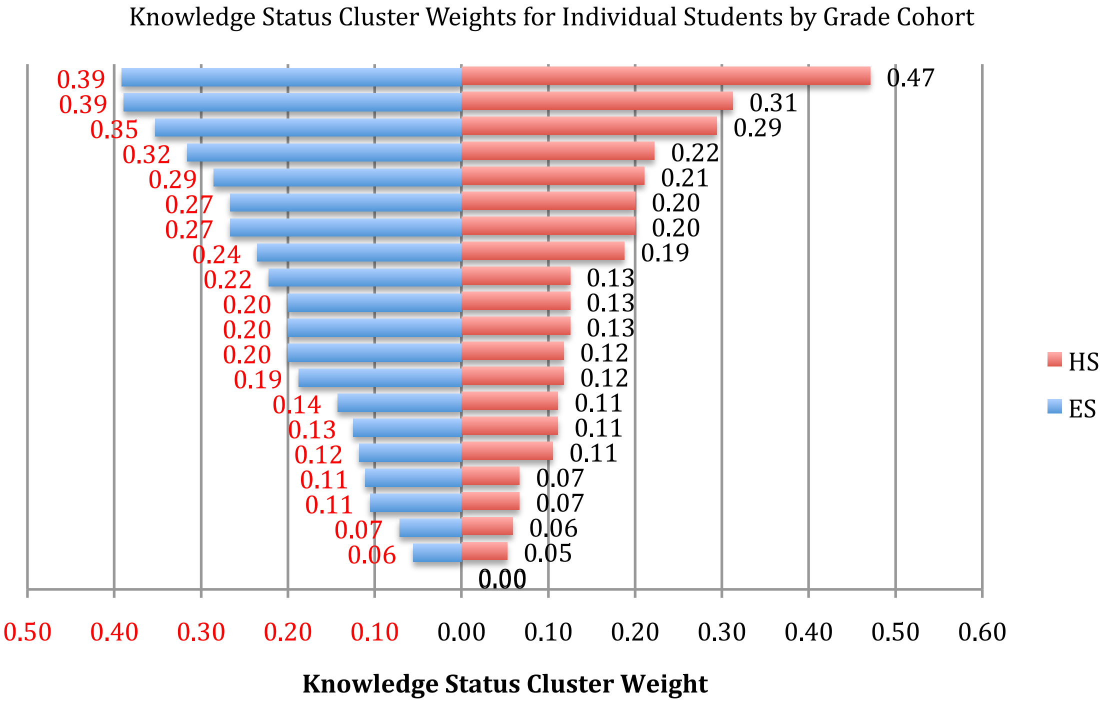
Table 11 shows the cluster weights for the elementary and high school grade cohorts. There were statistically significant differences between grade cohort cluster weights for the "Difficulty," "Interest" and "Contribution" clusters. Therefore, when making judgments about which scientific research questions are worthwhile to investigate, elementary school students considered difficulty and interest more so than did high school students. In contrast, high school students placed an even greater emphasis on the contribution that the research question will make.

Table 11: Aggregate Cluster Weights by Grade Cohort

	Knowledge Status	Difficulty	Interest	Contribution	Science Attributes	Non-Response
Aggregate Cluster Weights (ES)	0.20	0.09*	0.19*	0.43*	0.02	0.07
Aggregate Cluster Weights (HS)	0.16	0.05*	0.10*	0.65*	0.02	0.02
Aggregate Cluster Weights (All)	0.18	0.07	0.14	0.54	0.02	0.04

*significant (p<.05)

Despite the overall differences between grade cohort weights, it is important to note the variation in weights within each grade cohort. Figure 9 illustrates this variation by showing a distribution of individual student cluster weights for each grade cohort. In each cluster there is overlap between the ES and HS distributions. Visually, we can see that elementary school students consider difficulty and interest to a larger extent than do high school students. While there is one ES student who considered contribution to a large extent (large weight), on the whole, the high school students considered contribution to a greater extent to elementary school students.



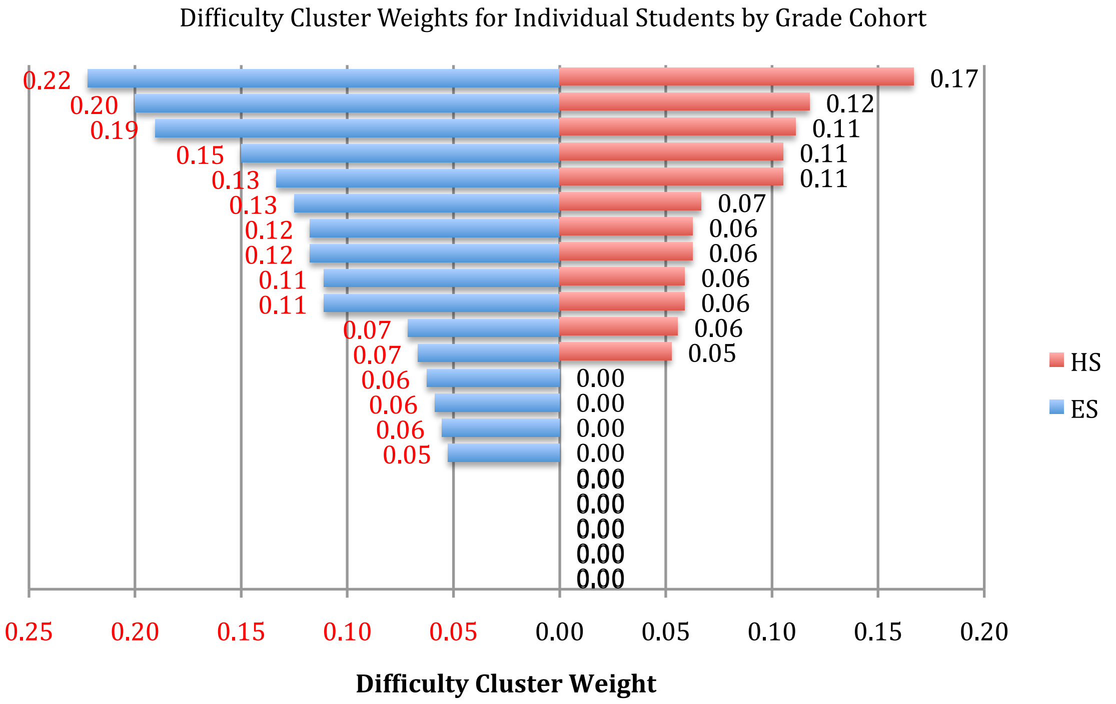
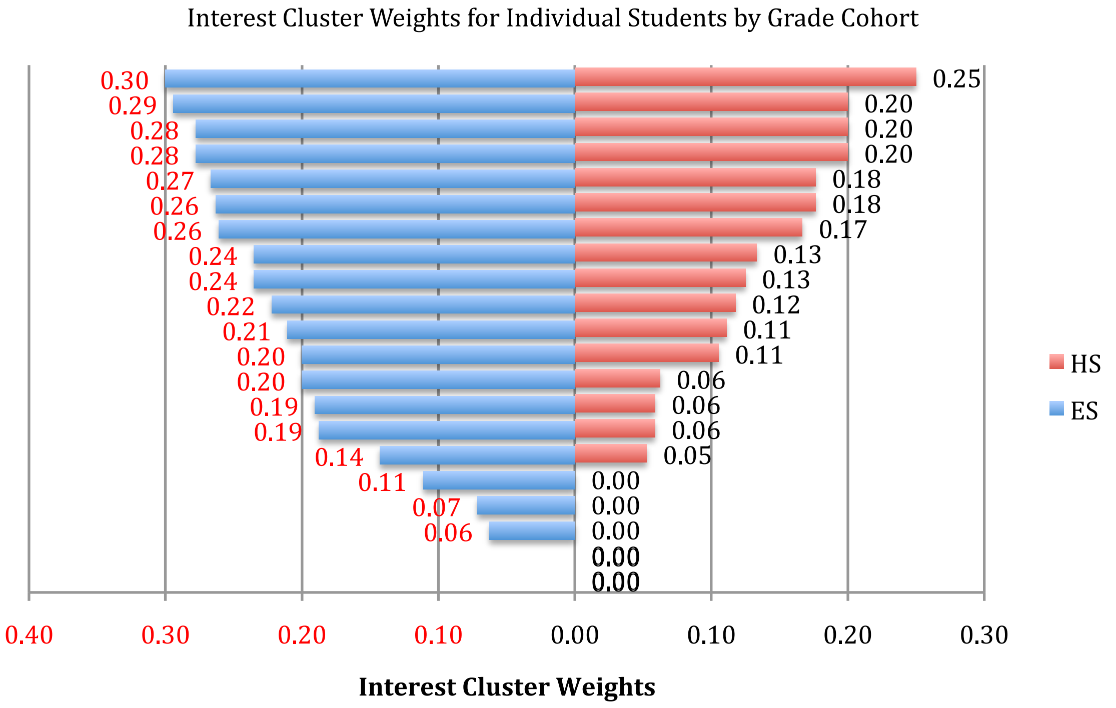


Figure 9: Distribution of Individual Cluster Weights by Grade Cohort



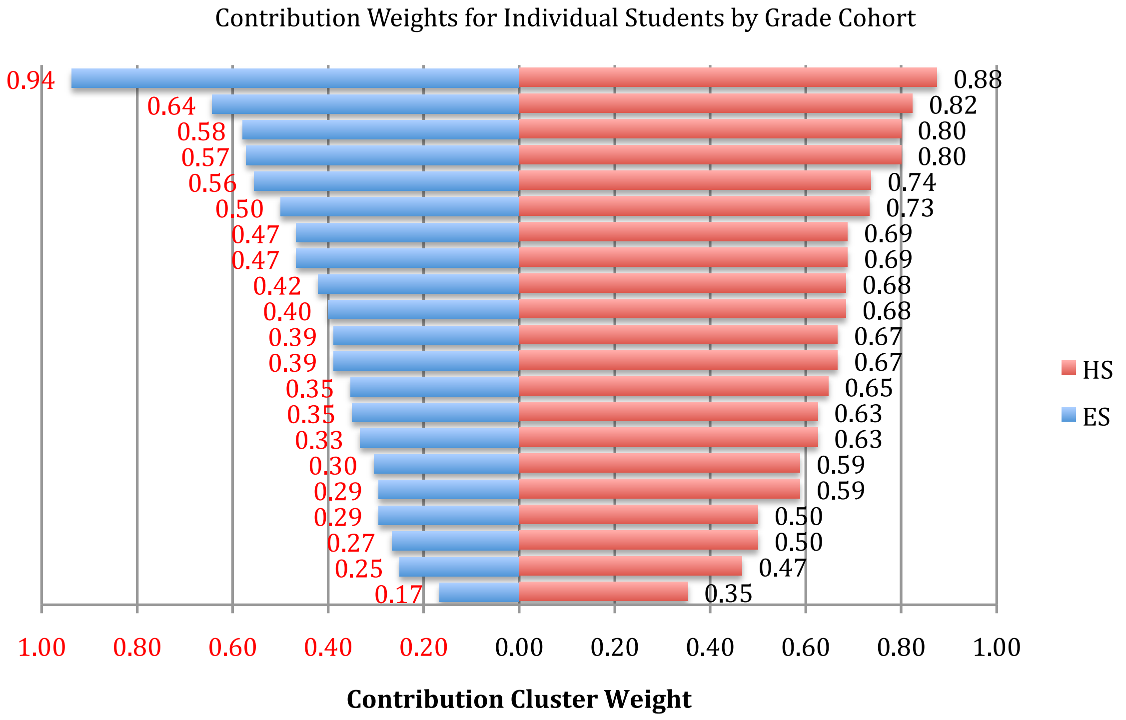


Figure 9: Distribution of Individual Cluster Weights by Grade Cohort (Continued)

There are several possible conjectures that could explain these findings. First, younger students have had less exposure to school science and thus may be relatively less aware of the kinds of contributions science has and can make to the world. Second, younger students may be more swayed by initial interest or curiosity rather than feeling the need for any further justification. Finally, the higher focus on contribution by high school students may be at least partially informed by the students' broader and more developed world-view. Older students are more aware of global or social problems, appreciate limited resources and realize the need to prioritize funding. Therefore older students are more likely to consider the specific contribution of the scientific research question. In Chapter 4, I unpack the "Contribution Cluster" and examine more specifically the nature of the contribution that students consider as well as patterns across grade cohorts and research questions.

To What Extent Do Students' Retrospective Criteria Reflect the On-line Criteria They Express While Making Their Judgments?

Thus far, the analyses have focused on the rationales that students expressed while explaining why they funded or did not fund each individual scientific research question. These rationales are vocalized "on-line" as students engaged in the decision-making process. This methodological tactic is used to gain insight into students' practical epistemology, the epistemic reasoning expressed in practice.

However, expressions of student rationales are not the same as their actual thought processes. Furthermore, students' explanations of their immediate decisions do not necessarily indicate that they possess a clear sense of the criteria they themselves are using. When asked to reflect on their decision-making process, to what extent do students enumerate the specific criteria that they used when making funding decisions?

To answer this question, I compare the funding decision rationales from the prior analysis to the criteria that students retrospectively report. During the Funding Allocation task, students provided "on-line" rationales for their funding decisions as they made their decisions. First they explained each decision not to fund and then they explained each decision to fund. At the end of the task, students were asked to reflect on their funding decisions, "How did you decide whether or not to fund the research questions? What was the difference between the questions you decided to fund and those that you decided not to fund?" These prompts ask students to articulate retrospectively the criteria they applied in their decisions.

When prompted, each student listed several criteria that *they thought* they used when deciding which scientific research questions were worthwhile. The criteria were coded using the same coding schema described in the prior section. However, each criterion was scored independent of the within-cluster coding constraints used in the prior analysis. This coding decision was made because each enumerated criterion was considered to be distinct in the mind of the student. So if they enumerated three criteria, each criterion received a code, regardless of the coding cluster. The following is an example of the retrospective criteria expressed by one student:

> I: *How did you decide which questions to fund and which questions not to fund?*
> S (HS_98): *I decided like based on how much it can help us. Like I feel like these ones (unfunded) wouldn't really help us that much and these (funded) would help the human population.*

In this case, the student reports that the sole criteria she used to make her funding decisions was the extent to which the research question would have a benefit for society.

Students' retrospective criteria used in funding decisions

To compare the "on-line" rationales with the "retrospective" criteria, I begin by examining the retrospective criteria. Table 12 shows how many students expressed retrospective criteria from each cluster. Of the 42 student subjects in this study, 28 students mentioned the contribution of the research question in their retrospective criteria. Furthermore, all but one high school student reported that they considered the contribution of the question in their funding decisions. Less than half of the elementary school students (eight of 21) mentioned the contribution of the research question in their retrospective criteria. However, more than twice as many ES students reported that they considered "Interest" as criteria in their decisions.

Table 12: Retrospective Criteria Counts by Cluster and Grade Cohort

Cohort	Knowledge Status*	Difficulty*	Interest*	Contribution*	Scientific Attributes	Total
ES	6	6	11	8	1	34
HS	4	4	5	20	2	35
ALL	10	10	16	28	3	69

*Note: Knowledge Status, Difficulty, Interest & Contribution ES vs. HS differences were statistically significant (p<.05)

Matching retrospective criteria with on-line funding rationales

Although each grade cohort reported using different criteria in its funding decisions, additional analysis is needed to determine how closely students' retrospective criteria matched their on-line criteria. To calculate the degree of matching, I start with the on-line cluster weights for each student (calculated in the prior section). Next, for each student, I sum the cluster weights only for the clusters that were *also* represented in their retrospective criteria.

For example, one student had the following on-line cluster weights: "Knowledge Status" (.32), "Difficulty" (0.0), "Interest" (.26), "Contribution" (.42) and "Scientific Attributes" (0.0). The student's retrospective criteria included "Knowledge Status" and "Interest" but did not include "Contribution," "Difficulty" or "Scientific Attributes." Therefore, the students' matching score is .58 [equal to the sum of "Knowledge Status" (.32) + "Interest" (.26)]. In other words, the students' retrospective criteria matched 58% of the students' on-line criteria. The larger the matching scores, the greater the overlap between the students' on-line and retrospective criteria.

Table 13 shows the mean matching scores for each grade cohort and all students. Overall, the students had a matching score of 0.56 across both grade cohorts. This means that their retrospective criteria captured 56% of their on-line cluster weights. However, high school students had a much higher matching score (0.74) than did elementary school students (0.38). This difference between ES and HS was statistically significant.

Table 13: On-line versus Retrospective Matching of Student Criteria

	On-line / Retrospective Matching Score	
	Mean	Range
All Students	0.56	0.00 to 1.00
ES	0.38	0.00 to 0.97
HS	0.74	0.47 to 1.00

Note: Difference between ES and HS matching scores is statistically significant (p<.05)

The large grade cohort gap in matching scores suggests that high school students were better able to generalize criteria that they had previously applied in decision-making. This is an important distinction. Talking about criteria and applying criteria are different, albeit related, cognitive practices. Criteria have meaning when they are applied and abstracting criteria from decisions is non-trivial. This distinction has implications for how epistemic reasoning is developed in classrooms. Discussing criteria divorced from their application may not sufficiently develop students' epistemic reasoning about "what counts" in science.

How Does Context (Professional vs. School) Affect These Judgments?

Even though the Funding Task targets students' *practical* epistemological reasoning through their decision-making, the task still focuses on worthwhile questions for *professional* scientists. In that sense, the task targets a hybrid of practical and formal epistemic reasoning. However, students' epistemic reasoning may vary depending on the context in which students are making epistemic decisions. Students' views of what counts in professional science may not be the same as their views of what counts in classroom science. This distinction has major implications for the image of science that is projected in schools.

Students' criteria in a science fair context

In school science, students engage in scientific practices and must decide what counts in that context. School science fairs are common formats for showcasing student-directed investigations. The criteria used to evaluate the quality of science fair projects function as a standard for what is valued in the discipline of science. The Science Fair Task is designed to elicit students' conceptions of what makes a scientific research question worthwhile in a school science context. In this task, students are asked to give instructions to science fair judges who are responsible for awarding the prize for "Best Scientific Research Question." Which criteria should judges use when deciding which research question is the *best* research question? Here is an example of a student response:

> I: *What instructions would you give the judges? How should they decide which research question is the best one?*
> S (ES_82): *Which one is more interesting, which one has a lot of details and which one talks about other stuff than just like saying the sun goes around the earth more than once a year and instead of just saying that maybe put more details into it.*

In this case, the student's comments indicate that he believes that a good science fair research question should be interesting and have a lot of details.

Student responses to the Science Fair Task were scored using the same process used to score the students' retrospective criteria in the Funding Allocation Task. However, there were criteria that students expressed in the Science Fair Task that were not expressed during the Funding Allocation Task. These criteria pertained to characteristics typically associated with schoolwork. Therefore, an additional coding cluster, "School Criteria," was also used to code the students' science fair criteria (Table 14). "School Criteria" include the formatting of the question (FOR), the details in the question (DTL) and the way that the question is presented by the student (PRE).

Table 14: Individual Codes for School Criteria Cluster

School Criteria Cluster

Code	Code Name	Description	Example(s)
FOR	Formatting	Student explains that science fair award should be based on formatting (i.e. style, penmanship, spelling, punctuation, etc.) of the research question.	"right capitals" "spelling correct" "neat handwriting"
DTL	Detail	Student explains that science fair award should be based on the extent to which the research question contains details or information.	"detailed" "more information in the question"
PRE	Presentation	Student explains that science fair award should be based on the quality of the presentation and articulation of the research question.	"if they don't remember it or concentrate on it,"

Students expressed a wide range of criteria that they thought should be considered when awarding a prize for "Best Scientific Research Question" at a science fair. Table 15 shows the counts of students who mentioned at least one science fair criterion in each of the different clusters. Roughly the same number of ES and HS students mentioned at least one criterion pertaining to "Knowledge Status," "Difficulty" and "Interest."

Table 15: Science Fair Criteria Counts by Cluster

	Knowledge Status	Difficulty	Interest	Contribution	Science Attributes	School Criteria
ES (N=21)	6	8	8	6	5	12
HS (N=21)	4	7	7	15	10	6
ALL	10	15	15	21	15	18

Note: ES vs. HS differences are not statistically significant ($p>.05$)

However, there were marked differences in the other criteria clusters. Roughly three quarters (15/21) of the HS students listed criteria relating to the contribution of the research question whereas only six of 21 ES students listed contribution criteria. Instead, more than half of ES students (twice the number of HS students) thought that the "Best Scientific Research Question" prize should be awarded based, at least in part, on school-related criteria. A similar pattern exists for HS students and "Science Attributes" criteria. These patterns may be explained by the difference in how schoolwork criteria are communicated to students as they advance through school. In elementary school, typical criteria for schoolwork may be more likely to include visual features like handwriting or spelling. However, as students advance through school and gain more experience with school science, the criteria begins to include attributes more specific to the domain (e.g., “Scientific” or “Testable”).

Comparing students’ professional and school science criteria

In this section, I compare the criteria that students used to evaluate scientific research questions in the school science fair and professional science contexts. In the Funding Allocation Task, students voiced retrospective criteria that they considered important for evaluating whether a scientific question is worthwhile for *professional scientists* to investigate. In the Science Fair Task, students listed criteria that they deemed important for evaluating if scientific research questions are worthwhile for *students* to investigate.

To determine how students’ criteria differ in the professional and school science contexts, I analyze the number of codes assigned to each student’s criteria in each context. To characterize the difference in criteria across contexts, I compare the count of codes in each context. Table 16 illustrates an example for one student. This student listed *one* “Interest Cluster” criterion in the professional context and *two* “Interest Cluster” criteria in the science fair context. Therefore, the student voiced an *increase* in “Interest” criteria in the science fair context (1 criteria -> 2 criteria.) The same analysis was conducted for each student in the study.

Based on the individual student changes in criteria between contexts, an aggregation of those changes provides a view of how student criteria changed across grade cohorts and overall. Table 17 shows counts of students and the change in their criteria between the professional (retrospective) context and the school (science fair) context. In the first three clusters (left side), we see that a little less than half of the students in each grade cohort changed the criteria that they used in the science fair context. However, the change was relatively evenly split between an increase and a decrease. There was a slight increase in high school students’ consideration of “Difficulty” and “Interest” in the science fair context.

Table 16: Change in Criteria Across Contexts for Sample Student

		Criteria Codes & Clusters					
	Student's Criteria	Knowledge Status	Difficulty	Interest	Contribution	Science Attributes	School Criteria
Professional Context (Retrospective Criteria)	Not really important				NIMP		
	Really fun activities for girls to do or boys and kids			INTS			
	You need something to live				SOC		
School Context (Science Fair Criteria)	Good writing on it like if you typed it out, right capitals						FOR
	Detailed question						DTL
	Worked really hard		EFF				
	Really interested in the question			INTS			
	Proud to do it, really honest			EMO			
Change in Code Counts		No Change (0->0)	Increase (0->1)	Increase (1->2)	Decrease (2->0)	No Change (0->0)	Increase (0->2)

Table 17: Changes in Student Criteria Counts (Professional vs. School) by Cluster

		Knowledge Status	Difficulty	Interest	Contribution	Science Attributes	School Criteria
ES Students	No Change	12	12	13	12	15	9*
	Increase	4	5	4	4	5	12*
	Decrease	5	4	4	5	1	0*
HS Students	No Change	17	12	12	7*	10*	15*
	Increase	2	6	6	4*	9*	6*
	Decrease	2	3	3	10*	1*	0*
All Students	No Change	29	24	25	19*	25*	24*
	Increase	6	11	10	8*	14*	18*
	Decrease	7	7	7	15*	2*	0*

Note: Increase/decrease refers to Science Fair criteria counts (i.e. an increase in the science fair criteria count compared to the retrospective criteria counts)
**(p<.05)*

More dramatic differences in students' criteria can be seen in the "Contribution," "Science Attribute" and "School Criteria" clusters. With the exception of elementary school "Contribution," these differences are statistically significant. The increase in "School Criteria" is not surprising, considering that the task is set in a school context. One possible explanation for the increase in "Science Attributes" is that students presumed that the science questions for professional scientists met the scientific criteria for research questions since professional scientists proposed the questions. However, in a school science fair context, the students were more likely to raise these criteria as an important standard for quality science fair questions.

Perhaps the most interesting result is the 48% (10 of 21) of high school students who mentioned fewer "Contribution" criteria in the school science fair context compared to the criteria they applied to questions for professional scientists. This finding suggests that high school students have different epistemic stances towards the importance of a scientific research question depending on who is investigating the question. One reason why a similar difference was not found in elementary students is because less than half of ES students considered contribution in their retrospective report of their own funding decisions (compared to 20 of 21 high school students).

It is important to note that the analytical approach used to compare professional and school epistemic criteria does not take into account the other differences between the Funding Allocation and Science Fair Tasks. The two tasks are not isomorphic and the different task designs mean that a controlled comparison of contexts is not possible. Nonetheless, students *are* enumerating criteria used to evaluate the worthwhileness of a scientific research question in both tasks and this provides legitimacy to the findings and analytic approach.

Summary

The findings in this chapter suggest that the epistemic reasoning involved in determining what makes a scientific research question worthwhile was not uniform among students. Given a set of candidate research questions, there was widespread disagreement about which of those questions were worthwhile to investigate. While certain questions in the study did have the highest funding rates and funding amounts, each question in the set was both supported and rejected by some students.

Analysis of the criteria that students applied in their judgments of worthwhile questions offers further insight into students' practical epistemologies. Students considered many different criteria in their reasoning about worthwhile research questions. These criteria included considerations of the knowledge status of the question, the difficulty of investigating the question, the degree of interest in the question, the contribution the question would make and the scientific attributes of the question. While younger students' and older students' decision-making reflected a range of shared criteria, the two grade cohorts differed in the weight they each placed on each criteria type. Furthermore, the criteria that students expressed during their "on-line" decision-making did not completely match the retrospective criteria they reported, suggesting that the articulation and application of criteria are not one in the same.

The contribution of a scientific research question featured prominently in students' epistemic reasoning. It was the highest weighted criteria cluster for both grade cohorts in their funding decisions. The contribution of the research question was also an area in which students' criteria changed the most due to context. For high school students in particular, the analysis identifies a diminished consideration of the research question's contribution in a school context relative to the professional scientist context. Taken as a whole, these findings warrant a deeper, more refined look at how students conceptualize the nature of a scientific contribution. In Chapter 4, I unpack the "Contribution" cluster to more closely examine the ways in which students think about a research question's contribution.

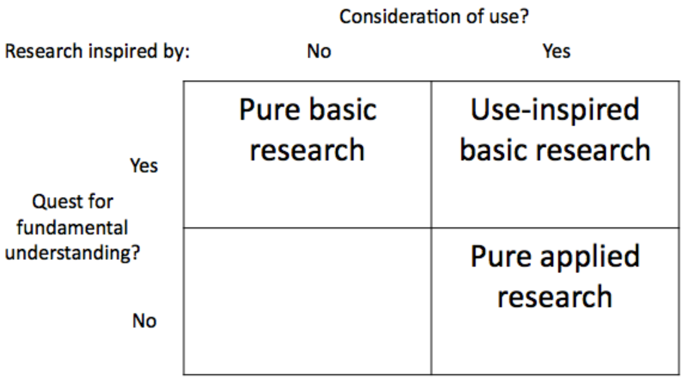
Chapter 4: How Do Students Conceptualize the Contribution of Scientific Research Questions?

In Chapter 3, I examined students' epistemic reasoning in the context of their judgments about which scientific research questions are most worthwhile. The analysis provided evidence that students consider several different criteria when evaluating scientific research questions. The most common consideration in students reasoning was the *contribution* of the research question.

In this chapter, I analyze the specific ways in which students conceptualized the contribution of a research question. Although students themselves did not use the term "contribution," students' rationales reflected consideration of the result, impact or consequence of investigating the research question.

In his seminal book on scientific research, *Pasteur's Quadrant*, Donald E. Stokes (1997) presents a useful framework for thinking about the contribution of scientific research. He traces how the goal structures of scientific research have been conceptualized, explaining that historically, the goal of scientific research has been viewed either as a "quest for understanding" (basic research) or with "consideration of use" (applied research). Research inspired by a quest for fundamental understanding seeks to advance our understanding of how the world works for the sake of expanding our knowledge. Research inspired by considerations of use is primarily concerned with addressing a specific problem in the world. This could include problems facing society or the planet or other pressing concerns.

Stokes (1997) argues that scientific research need not be dichotomized. Rather, scientific research can contribute simultaneously to understanding *and* use. He calls this "use-inspired basic research" and refers to that goal of research as "Pasteur's Quadrant" in deference to the major basic *and* applied contributions of Louis Pasteur (Figure 10).



Research inspired by:		Consideration of use? No	Consideration of use? Yes
Quest for fundamental understanding?	Yes	Pure basic research	Use-inspired basic research
	No		Pure applied research

Figure 10: Quadrant Model of Scientific Research (Stokes, 1997, p.73)

Stokes' model is useful as a framework for interpreting how students conceptualize the nature of the contribution of a scientific research question. To understand how students reason within Stokes' frame, I analyze patterns in students' responses to the Funding Allocation and Science Fair Tasks. These tasks are conducive to an analysis of students' considerations of the contribution of a scientific research question because the tasks do not presuppose any particular goal for science. Both tasks are framed in an open-ended way such that the students' criteria reflect what *they* believe to be the goals of science. I exclude student responses to the Ladybug Task because that task implies a purpose for investigating the scientific research questions that the students evaluate in the task.

In applying Stokes' framework, I reframe its dimensions so that they apply more directly to the evaluation of scientific research questions. To do this, I foreground the questions' *utility,* or what is gained by investigating the question. Thus, a question's "epistemic utility" reflects the extent to which the question is concerned with a quest for fundamental understanding (i.e., the utility of the question for advancing knowledge). Similarly, a question's "functional utility" reflects the extent to which the question is oriented around "considerations of use." The *functional utility* of a research question pertains to the utility of that question for solving a practical problem. Using the "utility" terminology, questions in Pasteur's quadrant have both epistemic *and* functional utility.

Coding schema for epistemic and functional utility

In order to operationalize how considerations of epistemic and functional utility are reflected in students' reasoning, I adapt the coding schema used in the prior chapter. Figure 11 shows which codes reflect epistemic and functional utilities. Students who reason that a question should be funded because it is Unknown (UK) or not funded because the question is Already Known (AK) are making a very basic judgment about the epistemic utility of a research question. In addition, rationales based on a Faulty Premise (FP) reflect a decision by the student that the question does not advance his or her understanding of the topic given that the question conflicts with something he or she already believes to be true. Later in the chapter, I provide examples of this reasoning to illustrate how this code reflects consideration of epistemic utility.

Student rationales coded as Personal Benefit (PRS), Social Benefit (SOC), Environmental Benefit (ENV), Doesn't Matter (DM) and Detrimental Outcome (DET) each reflect consideration of a question's functional utility. Each of these response types indicates that the student is evaluating the question in terms of the practical impact of investigating the research question on some type of problem.

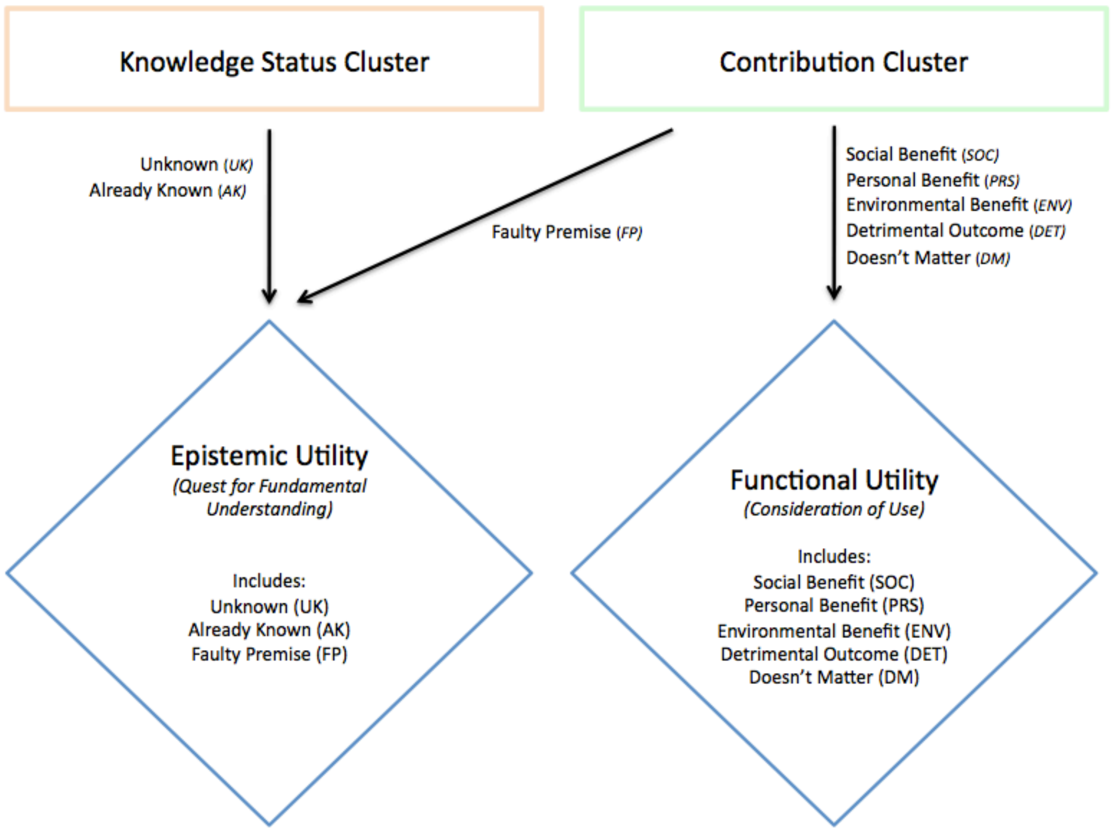


Figure 11: Coding Schema Aligned to Stokes' Framework

The analysis in this chapter is organized around the following questions:

1. To what extent do students consider the epistemic and functional utility of scientific research questions?
2. How do students reason about the epistemic utility of scientific research questions?
3. How do students reason about the functional utility of scientific research questions?
4. What factors mediate students' reasoning about the epistemic and functional utility of scientific research questions?

To address these research questions, I begin with an analysis of the prevalence of epistemic and functional criteria that students express in the Funding Allocation (on-line and retrospective) and Science Fair Tasks. Next, I turn to an analysis of specific patterns in students' criteria in relation to individual scientific research questions. This question-level analysis is important because, in the case of the Funding Allocation Task, students' evaluations about epistemic and functional utility are made in reference to the context of a specific research question. By analyzing patterns by question, we can understand how the particular question context informs the criteria that students use.

To What Extent Do Students Consider the Epistemic and Functional Utility of Scientific Research Questions?

The Funding Allocation and the Science Fair Tasks invited students to reason about what makes a research question worthwhile to pursue. Table 18 shows the proportion of students whose responses reflected considerations either of the epistemic utility or the functional utility, as well as students who considered both epistemic *and* functional utility in their responses to the tasks.

All students in the study considered the epistemic and functional utility of the candidate research questions as they decided which research questions to fund. However, this finding does not necessarily mean that students applied both types of criteria when evaluating each individual question, as would be suggested by Pasteur's Quadrant. When asked to list specific criteria for consideration (as in the retrospective prompt and Science Fair Task), very few students mentioned epistemic *and* functional criteria. Thus, while students do consider both utility types in the course of their reasoning, these criteria are not at the forefront of their minds, nor do they systematically use both when evaluating particular scientific research questions.

Table 18: Proportion of Students Considering Epistemic and Functional Utility

	Cohort	Epistemic Utility	Functional Utility	Epistemic & Functional Utility
Funding Rationales	ES	100%	100%	100%
	HS	100%	100%	100%
	ALL	100%	100%	100%
Funding Retrospective Criteria	ES	29%	29%	0%
	HS	19%	76%	5%
	ALL	24%	52%	2%
Science Fair Criteria	ES	29%	14%	5%
	HS	19%	48%	10%
	ALL	24%	31%	7%

Overall, students were more likely to list criteria reflecting functional utility than epistemic utility. Interestingly, in the retrospective and science fair contexts, elementary students were more likely to list criteria for epistemic utility and high school students were more likely to consider functional utility as a criterion for worthwhile research questions. To understand more deeply the specific ways that students considered epistemic and functional utility of research questions, I now turn to an analysis of patterns in students' reasoning across specific research questions.

How Do Students Reason About the Epistemic Utility of Scientific Research Questions?

In the Funding Allocation Task, students evaluated specific research questions one at time. Across the 14 candidate research questions, students had repeated opportunities to employ their tacit criteria about what makes a research question worthwhile as they decided whether or not to fund each research question. All students reasoned about the epistemic utility of a research question at least once during their on-line rationales. Students considered the epistemic utility of the research questions in relation to two rationale types: Knowledge Status (UK/AK) and Faulty Premise (FP). To understand how students applied these criteria, I analyze patterns for each form of reasoning.

Patterns in known vs. unknown rationales

Students' funding decisions involved students' own perceptions of the status of the knowledge targeted by each research question. This epistemic decision took two different forms. Students rejected the research question because the answer to the question was already known by them (AKS), by scientists (AKSC), by people in general (AKP) or because the information can be found on the internet or in books (AKIB). Alternatively, the student decided to fund the research question because it was unknown by them (UKS), by scientists (UKSC) or by people in general (UKP).

In cases when students reasoned that a question was Already Known (AK), the student was explaining why they decided *not* to fund the research question. It makes sense that if the question is already settled, investigating the question would not make an epistemic contribution. On the other hand, some students explained their decision to fund a research question based on the fact that the answer to the research question was still unknown (UK) and therefore worthwhile to investigate.

Table 19 shows the proportion of AK vs. UK rationales by grade cohort in students' on-line rationales. Overall, students were approximately three times more likely to justify their decision because the question was already known (AK) compared to unknown (UK). This trend was exaggerated even more so in the high school cohort.

Table 19: Proportion of UK vs. AK within Knowledge Status Cluster

	% of Codes (ES)	% of Codes (HS)	% of Codes (All)
Unknown (UK) %	33.3%	16.4%	26.2%
Already Known (AK) %	66.7%	83.6%	73.8%

While the ratio of UK to AK may be due, in part, to the specific set of research questions that students considered, the larger gap for high school students could be explained in two other ways. First, the high school students could be more sophisticated and less likely to consider the fact that a question is unknown as a sufficient reason to fund it. Second, the high school students may more commonly have justified their decisions on the basis that the question is already known because they are more aware of the existing scientific knowledge base. Either they themselves know more about the scientific world, or at least they know more about what is known by others (i.e. by scientists on in books, etc.)

Decisions about the epistemic contribution of a research question are made in relation to specific research questions. Table 20 lists the number of students who reasoned about the knowledge status for each research question. Students most commonly made epistemic judgments about the knowledge status of Q2-SODA, Q4-MOON and Q9-MAGNET. In each of those cases, students were more likely to reject those questions because they were "Already Known" (AK), a finding consistent with the broader patterns in the cluster. Furthermore, the high school students were more likely than the elementary school students to consider those questions as already known.

The two questions with the largest gap between grade cohorts were Q4 MOON and Q5-DINO. In the case of Q4-MOON, there were six more high school students than elementary school students who considered that question to be already known. This finding is likely due to the high school students' greater content knowledge about the moon. In the case of Q5-DINO, six elementary school students and no high school students reasoned that Q5-DINO should be funded because the color of dinosaurs' eyes is still unknown. This finding suggests that younger students are more likely to believe that increasing our knowledge about dinosaurs is worthwhile because a lot of information about dinosaurs is still unknown.

Upon closer analysis of the patterns in student's epistemic judgments of the knowledge status of scientific research questions, we find that students are more likely to reason that the question is already known than to justify a decision because a question is unknown. We also see that across different research questions, students apply the epistemic criteria of known vs. unknown unevenly. Certain research questions evoke this distinction more than others. Given the prevalence of the AK rationales over the UK rationales, it appears that students more commonly employ this basic epistemic distinction when they know the answer to the research question (or believe that someone else knows). Thus it may be easier to reject a research question on the basis of what one knows (as in the case of Q2-SODA and Q4-MOON) and even more so for high school students.

However, if students concede that the question is *not* already known, on what basis could they determine what epistemic contribution the question could make to our understanding of the world? In the next section, I examine this question more closely to determine the extent to which students' rationales may provide insight into how they reason about the specific epistemic contribution of a research question.

Table 20: Frequency of UK vs. AK by Question and Grade Cohort

Question	Knowledge Status	ALL	ES	HS	ES vs. HS Delta
Q1-TOWEL	UK	0	0	0	0
	AK	9	7	2	-5
Q2-SODA	UK	4	3	1	-2
	AK	17	7	10	3
Q3-METALS	UK	3	1	2	1
	AK	7	4	3	-1
Q4-MOON	UK	6	4	2	-2
	AK	18	6	12	6
Q5-DINO	UK	6	6	0	-6
	AK	4	3	1	-2
Q6-CHEMICAL	UK	4	3	1	-2
	AK	10	4	6	2
Q7-EARTHQUAKE	UK	1	1	0	-1
	AK	3	0	3	3
Q8-ALIENS	UK	0	0	0	0
	AK	3	3	0	-3
Q9-MAGNET	UK	4	2	2	0
	AK	12	7	5	-2
Q10-BEES	UK	4	3	1	-2
	AK	3	3	0	-3
Q11-MATH	UK	1	1	0	-1
	AK	5	4	1	-3
Q12-COMPUTERS	UK	0	0	0	0
	AK	5	4	1	-3
Q13-SOCCER	UK	1	1	0	-1
	AK	4	1	3	2
Q14-PESTS	UK	0	0	0	0
	AK	0	0	0	0

Patterns in "Faulty Premise" (FP) rationales

Research questions have epistemic utility if they advance the frontier of our understanding of phenomena. Therefore, evaluating the epistemic contribution of a research question implies a comparison of what we already know with what knowledge would be gained by investigating the research question. In order to fund their research, scientists must explain how their proposed research would build on and contribute to current research in the field.

While students are likely unfamiliar with the extent of scientific knowledge in the world today, they can still reason about the epistemic contribution of a research question *relative to their own understanding* of the world. By analyzing students' rationales about their funding decisions, we begin to see evidence of this type of epistemic reasoning.

Rationales coded as “Faulty Premise” (FP) (within the “Contribution Cluster”) are cases when student dismiss the research question by objecting to one or more of the premises underlying the research question. For example, a large number of students objected to the premise of the Q13-SOCCER research question: *Can boys or girls kick a soccer ball farther?* One elementary school student explained her decision not to fund this research question:

> S (ES_89): *It's just not, it's just like they can, they probably can kick it but it depends on what kind of person it is, like if they are a sports fan or if they play lots of sports a lot…well they just have to believe that they can.*

In this example, the student doesn’t believe that gender is the main determinant of how far someone can kick a soccer ball. Instead, the student argues that other factors are more important, such as how big a sports fan someone is, if they play sports a lot or if they believe in their own ability. Essentially, the student is comparing what she knows about the factors that determine how far someone can kick a ball with the gender focus of the proposed research question. She concludes that the research question does not productively contribute to our understanding of the phenomena of kicking a soccer ball since the focus of the question is on a variable (gender) that is not critical to determining how far someone can kick a ball. In this sense, by objecting to the basic premise of the research question (that gender matters), she reasons that the question does not have epistemic utility.

Approximately half of the “Faulty Premise” (FP) codes were assigned to students who had similar objections to the Q13-SOCCER question. However, there were other examples of students objecting to the premises of other research questions. Table 21 shows a sampling of other cases when students objected to the premise of the research question.

To be clear, in the case of “Faulty Premise” rationales, I am making an inference about the epistemic reasoning that students undertake when evaluating whether a question is worthwhile to investigate. Rejecting the premise of a question means that the assumptions of the question violate one’s own understanding. Scientists commonly investigate research questions to *test* conjectured theories and design specific experiments to rule out alternative explanations. Nonetheless, for scientists, each question has a specific epistemic utility for their research program. Questions that do not make epistemic contributions are not worthwhile for a scientist to investigate. While students’ reasoning in these tasks is not the same as the epistemic reasoning practiced by scientists, by rejecting the premise of a question, I argue that students are making judgments on an epistemic basis.

Table 21: Sample Student Faulty Premise (FP) Rationales and Interpretations

Research Question	Sample Student Response	Epistemic Interpretation
Q3-METALS	*(ES-65) All metals are kind of equal…*	Student objects to the premise that metals have different strengths. Therefore, investigating why some metals are stronger than others would not make an epistemic contribution since the question conflicts with her understanding of metals.
Q4-MOON	*(ES-86) The moon can't change shapes…*	Student objects to the premise that the moon can change shape. Therefore, investigating what causes the moon to change shape would not make an epistemic contribution since the approach conflicts with his understanding of the moon.
Q10-BEES	*(ES-77) I don't think they're really getting sick…*	Student objects to the premise that bees are getting sick. Therefore, investigating why bees are getting sick would not make an epistemic contribution since it is focused on a phenomenon that does not exist.
Q11-MATH	*(HS-111) Every kid has a different mind…*	Student objects to the premise that there is a best way to teach math since the student believes that every student learns math differently. Therefore, investigating the best way to teach kids math would not make an epistemic contribution since the approach conflicts with his understanding of how students learn.

Summary of findings of students' considerations of the epistemic utility of scientific research question

In this section I analyzed patterns in students' rationales in the context of epistemic utility of research questions. In evaluating scientific research questions, students made epistemic judgments about the knowledge status of the candidate research questions. Students more commonly rejected a research question because they considered the answer to the question to be already known, compared to funding a question because the answer is not yet known. This determination was, at least in part, dependent on the students' own content knowledge.

The "Faulty Premise" (FP) rationale can also be interpreted as reflecting epistemic reasoning about the contribution a research question would make to scientific knowledge (or to students' own understanding of the phenomena). I made the argument that by rejecting the premise of a question, the students are comparing what they know about the question to what the question is asking and thereby determining if the question would make an epistemic contribution.

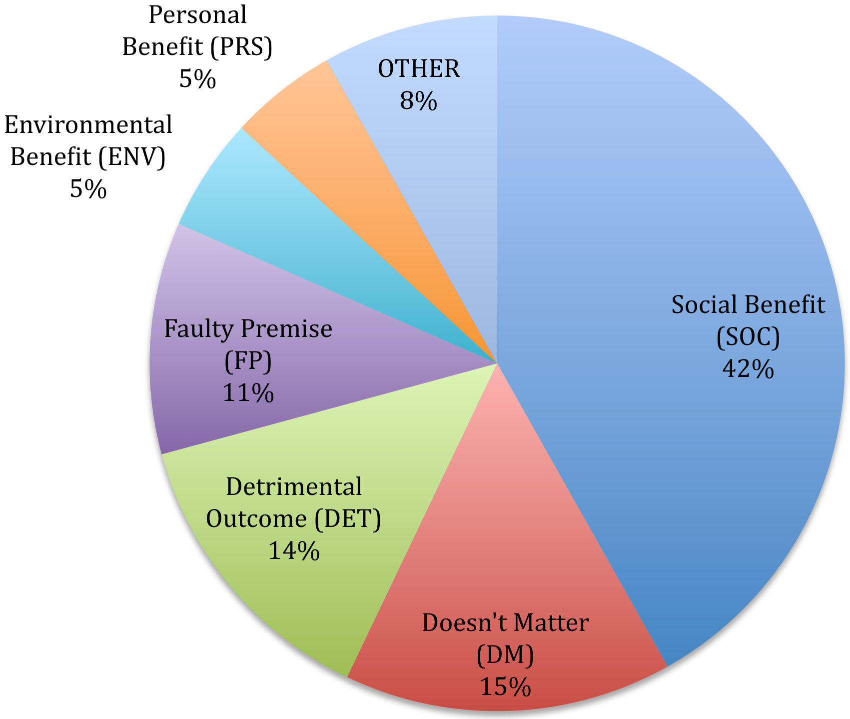
While the "Faulty Premise" lens may offer some insight into students' reasoning, there are limitations to the affordances of the Funding Allocation Task for deeply understanding how students reason about the epistemic contribution of a research question. Since the research questions in the Funding Allocation Task consist of a wide range of topics (from dinosaurs to earthquakes to metals), the task induces students to weigh topics against each other, in order to prioritize the distribution of funding. This may cause students to favor thinking in terms of what topics they think are most important from a practical standpoint, rather than evaluating the questions for their contribution to knowledge. In Chapter 5, I analyze results from the Ladybug Task, a task better suited to examining students' reasoning around epistemic utility of research questions. In the next section, I turn to an analysis of students' considerations of the functional utility of research questions in the Funding Allocation Task.

How Do Students Reason About the Functional Utility of Scientific Research Questions?

When students evaluate which scientific research questions are most worthwhile to investigate, a large part of their decision-making revolves around the contribution of the research question. In the prior section, I analyzed how student rationales reflected consideration of the question's *epistemic* contribution. In this section, I analyze how the students consider the *functional* contribution of the research question.

Aside from the "Faulty Premise" (FP) rationale (discussed in the prior section), the three most common "Contribution" rationales used by students to explain their funding decisions were "Social Benefit" (SOC), "Detrimental Outcome" (DET) and "Doesn't Matter" (DM) (see Figure 12). Taken together (including FP), these rationales constituted 82% of the "Contribution" cluster rationales. The SOC, DET and DM rationales can be reframed in a more general way to characterize students' evaluations of the research question as having a positive outcome (SOC), a negative outcome (DET) or having an outcome that doesn't matter at all (DM) and therefore not worthwhile.

Figure 12: Portion of Individual Codes within the Contribution Cluster



I now examine patterns in students' SOC, DET and DM rationales, identifying the research questions that were most commonly and least commonly evaluated on the basis of each rationale. I also examine the questions with the largest grade cohort gaps for each rationale.

Patterns in Social Benefit (SOC) rationales

The most common rationale that students used to explain the contribution of the research question was in terms of the beneficial outcome that the research would have for society. Table 22 shows the frequency of the "Social Benefit" (SOC) rationale as a percentage of "Contribution" codes and total codes by grade cohort. For high school students, the SOC rationale constituted a third of their overall rationales whereas SOC was the basis for only 11.9% of the funding decisions by elementary school students. This finding indicates that high school students were much more concerned with the benefit that the research question would have for society.

Table 22: Prevalence of Social Benefit (SOC) Rationale by Grade Cohort

Rationale Type	% of Contribution Codes (All)	% of Contribution Codes (ES)	% of Contribution Codes (HS)	% of All Codes (ALL)	% of All Codes (ES)	% of All Codes (HS)
Social Benefit (SOC)	41.8%	28.5%	50.7%	22.3%	11.9%	33.0%

Students invoked the SOC rationale for some research questions more so than for others. Table 23 shows the count of students who voiced the SOC rationale for each research question. Students most frequently cited the SOC rationale for research questions Q7-EARTHQUAKE, Q14-PESTS, Q11-MATH and Q10-BEES. Students' use of the SOC rationale for these questions was frequently motivated by their awareness of a particular social problem. For example, one elementary student's explanation for her decision to fund Q7-EARTHQUAKE indicates her awareness that earthquakes pose a real threat to people:

> S (ES_77): *Because in China there was this big earthquake and we need to know when they are coming so we have some place to put people where it doesn't run out of space.*
> I: *Why do we need to put people somewhere?*
> S (ES_77): *So they don't get hurt or broken bones.*

This student funded the Q7-EARTHQUAKE with $100 because she viewed the research question as having a major social benefit.

Table 23: Frequency of Social Benefit (SOC) Rationale by Question

Counts of Social Benefit (SOC) Rationale				
	ALL	ES	HS	Delta (HS-ES)
Q1-TOWEL	6	0	6	6
Q2-SODA	2	0	2	2
Q3-METALS	11	1	10	9
Q4-MOON	4	1	3	2
Q5-DINO	3	0	3	3
Q6-CHEMICALS	12	2	10	8
Q7-EARTHQUAKE	27	15	12	-3
Q8-ALIENS	13	2	11	9
Q9-MAGNET	5	0	5	5
Q10-BEES	18	4	14	10
Q11-MATH	19	6	13	7
Q12-COMPUTERS	15	5	10	5
Q13-SOCCER	2	0	2	2
Q14-PESTS	22	7	15	8
ALL	159	43	116	

One mediator of students' decision-making was the student's personal values. Students' values were reflected in their reasoning about which problems are important to investigate. For example, one high school student explained that finding alternatives to pesticides (Q14-PESTS) was a worthwhile research question based on the value he placed on the livelihood of organic farmers:

> S (HS_91): *It could be an advantage for the farmers, the organic farmers, to really help them gain again the fame that they used to have, selling products and not losing their farms and stuff.*

In another example, this elementary student's reasoning reflected a value placed on the importance of learning math, a value shared by many students:

> S (ES_65): *Math is my favorite subject and I would love to have other kids love math…math is actually the way you have to learn things in life pretty much everything you do in life has do to do with math.*

For some questions, there were large gaps in SOC rationales between elementary and high school cohorts (see Table 23). The largest gaps were for Q10-BEES, Q3-METALS, Q8-ALIENS, Q6-CHEMICALS and Q14-PESTS. In all these cases, the high school students voiced a SOC rationale more often than did the elementary school students.

One explanation for these grade cohort gaps is that high school students were more aware of the problems these questions sought to address. For example, the Q10-BEES question asks, "Why are bees getting sick?" High school students were more likely to know about bees' role in pollination and more likely to make the connection between bees getting sick and the potential impact on the environment, food system and ultimately on people. One high school student explained this concern:

> S (HS_99): *Bees produce food and help spread pollen that produces food, if they die or get sick then that means everything starts to fail, cause it has a big part in our world…there's not enough food, so we don't need any more stuff going away. Plus they have honey, so that's good.*

If elementary school students are less knowledgeable about bees and their role in the food system, they will be less likely to view bee sickness as a problem. Of the four elementary students whose rationales were coded as SOC, three students expressed concern over the risk to honey and only one student made the connection to pollination and the ecosystem. Thus student considerations of the social benefit (SOC) likely depended on whether they viewed the research question as addressing an important social problem. The students' values, awareness of the problem and prior knowledge about the problem are all factors that influence their decisions about which questions are most worthwhile to investigate.

Patterns in “Doesn’t Matter” (DM) rationales

In contrast to the rationales based on “Social Benefit” (SOC), many students decided that a question was not worthwhile because it was pointless or “Didn’t Matter” (DM). Table 24 shows the frequency of the “Doesn’t Matter” (DM) rationale as a percentage of “Contribution” codes and total codes by grade cohort. For high school students, the DM rationale constituted twice as large a proportion of their overall rationales (8.5%) compared to the elementary school students (3.9%). These findings indicate that the high schools students more readily determined that an issue or problem was trivial.

Table 24: Prevalence of Doesn’t Matter (DM) Rationale by Grade Cohort

Rationale Type	% of Contribution Codes (All)	% of Contribution Codes (ES)	% of Contribution Codes (HS)	% of All Codes (ALL)	% of All Codes (ES)	% of All Codes (HS)
Doesn’t Matter (DM)	15.3%	11.3%	17.9%	9.9%	5.8%	13.9%

As with the SOC rationale, students invoked the DM rationale for some research questions more so than for others. Table 25 shows the count of students who voiced the DM rationale for each research question. Students most frequently cited the DM rationale for research questions Q1-TOWEL and Q5-DINO. There was a large gap between grade cohorts for these two questions and for Q12-COMPUTERS. In the case of Q12-COMPUTERS, the high school students thought that computers are already sufficiently fast and therefore the research question was pointless.

Table 25: Frequency of Doesn’t Matter (DM) Rationale by Question

Counts of Doesn’t Matter (DM) Rationale				
	ALL	ES	HS	Delta (HS-ES)
Q1-TOWEL	15	4	11	7
Q2-SODA	2	0	2	2
Q3-METALS	2	2	0	-2
Q4-MOON	2	0	2	2
Q5-DINO	13	4	9	5
Q6-CHEMICALS	0	0	0	0
Q7-EARTHQUAKE	0	0	0	0
Q8-ALIENS	3	1	2	1
Q9-MAGNET	3	0	3	3
Q10-BEES	3	1	2	1
Q11-MATH	3	3	0	-3
Q12-COMPUTERS	8	1	7	6
Q13-SOCCER	4	1	3	2
Q14-PESTS	0	0	0	0
ALL	58	17	41	

Roughly half of the high school students thought that the Q1-TOWEL and Q5-DINO research questions did not matter (DM). The students characterized these questions as “pointless.” One student explained her reasoning about why she discarded Q1-TOWEL:

> S (HS_99): *Because it makes...it's not going to be productive to the world, like it does not matter which one is best. It's just doing something, it's not going to be a natural thing to help the world.*

Another student explained why she did not fund the question about the color of dinosaur eyes (Q5-DINO):

> S (HS_100): *I feel like that's not going to get us anywhere sort of…it's just like a random fact.*

In these examples, the students discounted the merit of the research question because they did not think that the question would make a contribution to the world. While the elementary school students were more likely to fund Q5-DINO because it was interesting and they liked dinosaurs, the high school students were more likely to view the question as a “random fact,” and therefore not worthwhile.

Curiously, there were several elementary school students who thought that Q11-MATH (What is the best way to teach kids math?) didn’t matter (DM). Those students did not view math instruction as a problem. They thought that the way kids currently learn math is just fine:

> S (ES_66): *Kids from today are learning math really well and they don’t need new ways to teach math….*
>
> S (ES_89): *They say that they can teach them math…they say that they can teach them math but usually they get it on their own…they know it, they know the answer they just can't explain it in a good way….*

These are additional examples of cases where students who are not aware of a problem (i.e., that some kids struggle to learn math) are unlikely to view a scientific research question about that problem as worthwhile.

Patterns in “Detrimental Outcome” (DET) rationales

Some students decided not to fund a scientific research questions if they thought investigating the research question would lead to an undesirable or detrimental outcome. These cases were coded as “Detrimental Outcome” (DET). Table 26 shows the prevalence of DET rationales as a percentage of “Contribution” codes and of total codes by grade cohort. While DET codes comprised twice the proportion of “Contribution” codes for elementary school students compared to high school student, the proportion of DET codes overall was roughly the same for each cohort.

Table 26: Prevalence of Detrimental (DET) Rationale by Grade Cohort

Rationale Type	% of Contribution Codes (All)	% of Contribution Codes (ES)	% of Contribution Codes (HS)	% of All Codes (ALL)	% of All Codes (ES)	% of All Codes (HS)
Detrimental Outcome (DET)	13.7%	19.9%	9.6%	7.3%	8.3%	6.3%

Students considered some scientific research questions to be more detrimental than others. Table 27 shows the count of students who voiced the DET rationale for each research question. Q2-SODA (What happens when plants are given soda instead of water?) was the research question that the most students viewed would have a detrimental outcome (DET). The following elementary school student's response typifies student thinking about this question:

> S (ES_63): *That's just an automatic no…soda is bad for you that I don't even know anything that is badder than soda….*

This student is rejecting the scientific research question on the basis of her prior content knowledge. She knows that soda is unhealthy for people and reasons that the soda would also be unhealthy for plants. While there may indeed be benefits to giving soda to plants, this student is using her prior knowledge to anticipate a negative outcome, deeming the research question not worthwhile.

Table 27: Frequency of Detrimental Outcome (DET) Rationale by Question

Counts of Detrimental Outcome (DET) Rationale				
	ALL	ES	HS	Delta (HS-ES)
Q1-TOWEL	3	2	1	-1
Q2-SODA	18	10	8	-2
Q3-METALS	1	1	0	-1
Q4-MOON	0	0	0	0
Q5-DINO	2	1	1	0
Q6-CHEMICALS	5	4	1	-3
Q7-EARTHQUAKE	2	1	1	0
Q8-ALIENS	1	0	1	1
Q9-MAGNET	2	1	1	0
Q10-BEES	4	4	0	-4
Q11-MATH	0	0	0	0
Q12-COMPUTERS	5	3	2	-1
Q13-SOCCER	9	3	6	3
Q14-PESTS	0	0	0	0
ALL	52	30	22	

As was evident in the analysis of the previous code patterns, students' personal values also influenced their assessments of which research questions would lead to detrimental outcomes. The Q13-SOCCER (Can boys or girls kick a soccer ball farther?) evoked many concerns about the outcome of investigating that research question. One student explains her reasoning:

> S (HS_92): *Not that one cause yeah... cause I've seen some girls kick really far soccer balls, but I've seen guys do it too. I mean I feel like if you really ran an experiment about that and then like one sex happened to be larger than the other…that might start something larger than what it really needs to be. And it might become something that's more towards sexist than the actual experiment and scientific data.*
> *I: When you say it might start something larger, what do you mean by that?*
> S (HS_92): *Because I remember arguing with people in elementary about guys who were better than girls and girls are better than guys so, when it comes to like kicking a soccer farther I feel like if they find out that one sex can do something more than the other sex, they might put that towards other things. Like well what can other sexes do at sports or other sexes, yeah. That might try to have like a domino effect towards other things. Where it's coming towards data where they can have actual scientific data that says well girls are better than guys and guys are better than girls.*
> *I: And so then what would happen? I want to make sure I understand what you're thinking.*
> S (HS_92): *I don't want it to become something that's sexist. I want it to stay towards scientific data.*

It is clear from these examples that students' personal values around gender equality and anti-sexism played a large role in their judgments about the potentially detrimental impacts of investigating Q13-SOCCER. In this way, personal values can dictate judgments about whether a scientific research question is worthwhile to investigate or not.

Personal values played a larger role in high school students' judgments about the detrimental impacts (DET) of research questions compared to elementary school students. Many of the elementary school students' DET rationales pertained to the perceived risks in actually carrying out the investigation. For example, some elementary schools students were concerned that the bee might sting (Q10-BEES) or that the chemicals might endanger the scientists (Q6-CHEMICALS). These concerns led students to reject the research questions on the basis of the potential for a detrimental outcome (DET). Thus students' judgments about the detrimental outcome of scientific research questions can be motivated by their prior content knowledge (e.g., Q2-SODA), personal values (e.g., Q13-SOCCER) or the potential danger in conducting the investigation itself (e.g., Q6-CHEMICALS.)

Further evidence of the influence of students' personal values

As evidenced by analysis in the prior section, students' personal values influenced their judgments about the potentially detrimental outcomes that could result from investigating a scientific research question. Personal values were also reflected in other types of students' rationales.

One area of importance for students was protecting the environment. Students' environmental values were reflected in students' rationales for their funding decisions. The following student expressed her environmental values in the context of space exploration (Q8-ALIENS: Are there living things on other planets?)

> S (HS_103): *I think we need to figure out how our own planet works before we start look for other living things on other planets…if they're putting a lot of energy into finding out if there's a possibility of life that's money taken away from figuring out how we can sustain our own environment.*

Another student expressed his concern about technology's impact on the environment in his explanation for why he did not fund Q12-COMPUTERS (How can we make computers run faster?):

> S (HS_93): *It's just going to waste money on technology when they could use it for the environment or something; satellite and all that stuff, it's messing up our environment….*

These are several examples of how students' environmental values influenced their funding decisions. Debates about how funding should be spent in the political sphere are often driven by people's value systems as well. In this case, we find that students' personal values also impact how they think science funding should be distributed.

Individuals' religious beliefs can also play a prominent role in how they view scientific knowledge and future directions for science, as evidenced by the debates about evolutionary theory or stem cell research. In this study, there were several students for whom religious views explicitly influenced their decisions about worthwhile scientific research questions. These students did not fund one of the most popular research questions, Q7-EARTHQUAKE (Is it possible to predict when an earthquake will happen?):

> S (ES_90): *Cause nobody can find if an earthquake is about to come. Well I mean god knows.*
>
> S (HS_93): *No. I don't know but just my religion says you can't predict this kind of stuff, so...in my religion, it's not true.*

In both of these examples (coded as NPOS "not possible"), students decided not to fund the Q7-EARTHQUAKE research question on the basis of their religious beliefs.

This lends further evidence that one's individual beliefs can also impact epistemic decisions.

Summary of findings of students' considerations of the functional utility of scientific research questions

In this section I analyzed patterns in students' rationales in the context of the functional utility of research questions. Students evaluated the contribution of the research question in terms of the impact of the question on a problem facing society. Most commonly, students considered the outcome of the research question as having a social benefit (SOC), not mattering at all (DM) or having a detrimental outcome (DET).

Students generally considered a research question to be worthwhile if the question addressed a problem that they considered important. However, there were different factors that mediated students' judgments about worthwhile problems. First, students needed to view the research question as addressing a real-world problem. In some cases, students were unaware that an issue was a problem due to their own prior knowledge about the issue. For example, elementary school students were less likely to appreciate the role of that bees play in our ecosystem (Q10-BEES) or appreciate that there are many students who struggle to learn math (Q11-MATH). In those cases, students were unlikely to fund a research question that addressed an issue that they did not view as a problem.

Students' personal values and beliefs also mediated their judgments about which problems were worthwhile to address. Students chose to fund research questions that they thought made a contribution to social problems aligned with their value systems. In some cases, students' values also dictated how they viewed the outcome of the investigation. For example, many students' values around gender equality influence their decision not to fund the Q13-SOCCER research question. Furthermore, students' religious belief systems also played a role in their judgments about the *worthwhileness* of scientific research questions (e.g., Q7-EARTHQUAKE).

Chapter Summary

The analyses in this chapter indicate that students considered both epistemic and functional utility in their reasoning about what makes a scientific research question worthwhile to investigate. On an epistemic basis, students considered whether or not the research question was already known. Some students also reasoned about the specific epistemic contribution on the basis of a comparison with their prior knowledge of the subject. On a functional basis, students who viewed the question as addressing a problem that they viewed as important and aligned with their personal values were more likely to consider a research question to be worthwhile.

Empirical findings indicate that students more commonly considered the functional contribution of the research question than they did the epistemic contribution. This finding is not surprising since the students were positioned to allocate funding across such a wide range of questions. By design, the Funding Allocation Task's diversity of research topics and limited available funding forced students to make tough choices, and to prioritize which problems they viewed as most important to address. With only one research question in any given topic (e.g., only one question about chemicals), students were not being asked to make fine-grain, epistemic distinctions between more than one research question in the same topic.

In evaluating the *epistemic utility* of research question, one needs to evaluate the contribution to understanding relative to what is already known. This determination depends on having sufficient conceptual understanding to know if a new question is actually advancing our current knowledge. In the next chapter, I turn to an analysis of student responses to a different task, one that is specifically designed to more adequately and deeply investigate students' reasoning about the *epistemic utility* of scientific research questions.

Chapter 5: How Do Students Evaluate and Coordinate a Set of Scientific Research Questions to Investigate an Idea?

This chapter focuses on students' evaluations of the epistemic utility of scientific research questions. The analysis in this chapter complements the approach taken in the prior analytical chapters. In Chapters 3 and 4, I examined how students' practical epistemologies were reflected in their rationales about what makes a scientific research question worthwhile to investigate. The analysis revealed that a research questions' expected contribution is a major consideration for students in determining if that question was worthwhile.

However, the specific contribution of a research question was conceptualized in many different ways. In Chapter 4 examined evidence indicating that students conceptualized the contribution of a research question largely in terms of the social or environmental benefit that investigating the question might have. This finding maps to Stokes' (1997) notion of "consideration of use."

Yet even if scientific research questions can and often do have benefits for society, scientific research questions must also have epistemic utility. In fact, if science is a knowledge-building enterprise, then the epistemic utility of research questions is unequivocally central to what makes a research question worthwhile. In essence, worthwhile scientific research questions make contributions to our understanding as part of a knowledge-building process. Therefore, considering the epistemic utility of a scientific research question is of critical importance. If a scientific research question does not have epistemic utility, then it does not contribute to our knowledge and therefore may not be worthwhile to investigate.

To better understand how students reason about the epistemic utility of scientific research questions, I examine student responses to a task designed to target this type of reasoning. The "Ladybug Task" is a novel task designed to elicit students' ideas about worthwhile research questions in a particular scenario with a goal structure that involves testing an idea. By evaluating the questions in relation to their utility for testing an idea, we can examine how students reason about the epistemic utility of the research questions.

The Ladybug Task is situated in a particular ecological context. A primary challenge for scientists who study agroecology is to find sustainable methods for managing crop pests. One method that scientists research is a technique called "floral resource provisioning." In this method, flowers are interspersed among crop plants in order to create resources and habitat for beneficial insects that can help control the pest population.

There are two primary theories that have been advanced by researchers in this field to explain the relationship between floral resource provisioning and pest reduction. The first theory is the natural enemy theory, which posits that the flowers attract natural enemies of the pests because the enemies are attracted to the carbohydrates found in pollen and nectar. The natural enemies then feed on the pests for protein and thus serve to control the pest population. An alternative theory is the resource concentration theory. This theory posits that the diminished pest colonization results from the camouflaging of the crop by the flowers. The presence of flowers dilutes the target crop, thereby increasing the chance that the pest will land on the flower instead of the broccoli or that the pest will be chemically repelled by the flower. It is also possible that a combination of these two theories (natural enemies and resource concentration) best accounts for the mechanism. The Ladybug Task is designed based on a simplified model of an ecological system and scientific research program. Therefore, a design decision was made to focus primarily on the "natural enemies" hypothesis.

In the Ladybug Task, students were asked to evaluate a set of five candidate research questions posed by a hypothetical student. Students decided which questions they thought made sense to investigate. Then students explained *why* they thought the questions either did or did not make sense to investigate. Finally, students were also asked to explain in what order they would investigate the questions. Students also provided a rationale for *why* they thought the questions should be investigated in that particular order.

In this chapter I analyze patterns in student responses in several ways. I begin with an analysis of the percentage of students who selected each question as worthwhile, disaggregated by grade cohort. Next, I analyze the different types of rationales that students provided for their decisions, with particular attention to the epistemic utility that students did or did not attribute to the questions. I then turn to an analysis of the ways in which students sequenced the research questions for investigation and the rationales students gave for the question order as a reflection of their consideration of the epistemic contingency between questions. The goal of the analyses is to understand how students in different grade cohorts reason about the utility of scientific research questions and how these reasoning patterns compare to normative scientific notions about the epistemic utility of these questions.

As a backdrop to the analysis in this chapter, I consulted with an expert scientist in the field of agroecology who conducts research within the same ecological context as the one used in the Ladybug Task. The expert scientist provided his own assessments of the epistemic utility of the research questions in this task. While the expert's judgments are by no means definitive, they do provide a productive frame of reference within which to interpret student responses to the task. In this sense, his view is taken as normative for the purpose of comparison with student responses.

The analysis in this chapter is organized around the following questions:

1. Which questions do students think make sense to investigate?
2. What utility do students attribute to different scientific research questions?
3. How do students sequence research questions to investigate an idea?
4. On what basis do students rationalize the coordination of research questions?

Which Questions Do Students Think Make Sense to Investigate?

In the Ladybug Task, the hypothetical student wants to conduct a science investigation so that he or she can test whether the presence of flowers will help manage a pest population on a field of broccoli plants. The hypothetical student suggests five candidate questions to investigate (Table 28). Students in this study were asked to decide which of the five questions they thought made sense to investigate if the hypothetical student wanted to test his/her theory. Students could select more than one research question if they thought that more than one question made sense to investigate.

Table 28: Candidate Research Questions for Ladybug Task

Question #	Handle	Research Question
Q1	HowMany	How many aphids can a ladybug eat in one day?
Q2	NextTo	Does planting flowers next to broccoli cause fewer aphids?
Q3	Spots	How many spots do ladybugs have on their backs?
Q4	WhichFlowers	Which flowers do ladybugs like best?
Q5	EatFirst	Which part of broccoli plants do aphids eat first?

According to the expert scientist, Q1-HowMany, Q2-NextTo and Q4-WhichFlowers make sense to investigate. Each has a particular epistemic utility for investigating the phenomena. Investigating Q1-HowMany provides information about the ladybugs' appetite that is valuable for understanding the nature of feeding relations between the two organisms and to determine if ladybugs eat sufficient numbers of aphids to determine if the ladybug strategy is feasible. Q4-WhichFlowers is useful for understanding the nature of the relationship between ladybugs and flowers. Ladybugs may have different flower preferences and testing the effectiveness of planting flowers depends on a clear understanding of these preferences. Finally, Q2-NextTo is a question that tests the main hypothesis of the experiment. In the final section of this chapter, I will elaborate on the particular ways that the expert scientist recommends investigating these questions.

Following the scientist's normative perspective on the utility of these questions, I now turn to students' judgments about which questions make sense to investigate. Figure 13 shows the percentage of students who thought that each question made sense to investigate. Three questions (Q1-HowMany, Q2-NextTo and Q4-WhichFlowers) were selected by the highest percentage of students. There were minor differences between grade cohorts for these questions but overall, the vast majority of both cohorts agreed that these questions would make sense to investigate. No high school students and only 14% of elementary students thought that Q3-Spots made sense to investigate. Half of the students thought that Q5-EatFirst made sense to investigate but there was a large disparity in selection rates between elementary (74%) and high school students (29%) for that question.

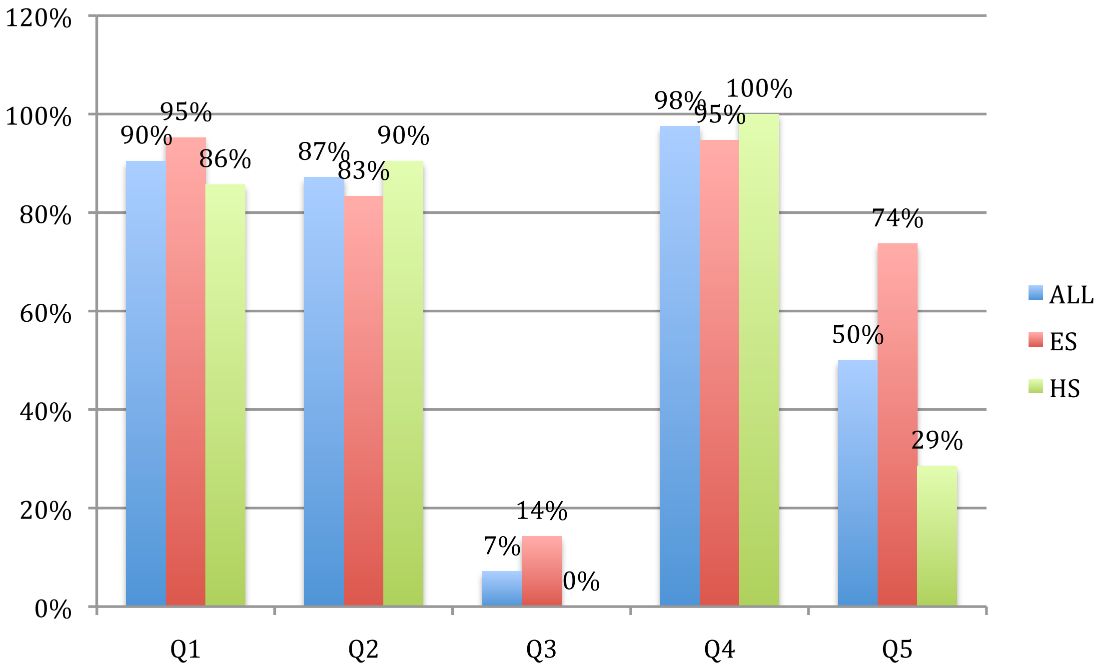


Figure 13: Percentage of Students Selecting Each Question in Ladybug Task

Note: ES vs. HS differences in selection percentage are not statistically significant (p<.05)

With the exception of Q5-EatFirst, there was widespread agreement among the students about which questions made sense to investigate and which did not. In two cases, Q3-Spots and Q4-WhichFlowers, there was complete agreement among high school students. In the other cases, there was a minority of students in each cohort who did not agree with the majority. But overall selection rates of 90%, 87% and 98% for Q1-HowMany, Q2-NextTo and Q4-WhichFlowers respectively and 7% for Q3-Spots indicate widespread agreement. This finding suggests that in this task, at least on the surface, students mostly agreed on which questions make sense to investigate. However, in the next section I will explore the reasons students gave for why a particular question makes sense to investigate. This will also allow for further investigation into the disparities observed for Q5-EatFirst.

On What Basis Do Students Make Judgments About Which Scientific Research Questions Make Sense to Investigate?

In the prior section, findings suggested that students largely agreed about which questions made sense to investigate. However, to better understand the extent to which students based their judgments in considerations of epistemic utility, I now turn to an analysis of the students' rationales for their decisions. I begin by presenting a coding scheme to interpret students' rationales. I then analyze patterns in the rationales that students gave for their judgments of each of the five scientific research questions. I attend to patterns within and across grade cohorts.

Coding schema

Similar to the analytical approach used in Chapter 4, I developed an emic coding schema through an iterative process. In a preliminary review of student rationales, I identified instances of reoccurring rationales and assigned codes to each rationale type that occurred more than one time. As the schema emerged, I applied the schema to the data making adjustments as necessary so that the schema adequately captured the rationales stated by students.

The individual codes are organized into clusters as shown in Table 29. There are five primary coding clusters. The first cluster pertains to student rationales that attribute epistemic utility to the research question that has implications for the research program as a whole. This means that the student is thinking in terms of how the question functions to organize and coordinate the other research questions as part of a connected research program.

The second cluster pertains to rationales that consider the utility of specific knowledge that would be gained by investigating that question. This cluster includes a wide range of codes that reflect students' considerations of epistemic utility, the usefulness and importance of particular information resulting from the investigation.

The third cluster pertains to student rationales in which they do not consider the research question to be relevant to testing the idea. In these cases, the students judge that the question does not have epistemic utility.

The fourth cluster pertains to student rationales that consider utility, however the utility is not epistemic in nature. For example, they view the question as useful because it will have a generally beneficial outcome or will make the experiment work, but they do not consider the utility in terms of the knowledge or information that the question will provide.

The fifth cluster pertains to student rationales that either do not consider the utility of the question or their consideration of utility is very weak. For example, some students rationalized their decisions based on the difficulty of the question to investigate [EA] or because it was interesting [INTS]. In other cases, the students did view the question as having utility, but a utility simply for their own personal benefit [PRS] unrelated to the theory being tested or to allow them to tell other people [TTO]. The final two clusters refer to cases in which students either state that they do not have a reason for their decision or they were not explicitly asked for a rationale for a particular question [NA] (missing data).

Table 29: Individual Codes for Student Rationales by Cluster (Ladybug Task)

Program Utility Cluster

Code	Code Name	Description
FEA	Feasible	Student believes the question makes sense to investigate because the question helps determine the feasibility of the idea that the hypothetical student wants to test.
MAIN	Main Question	Student believes the question makes sense to investigate because the question gets at the main idea that the hypothetical student wants to test.

Specific Epistemic Utility Cluster

Code	Code Name	Description
HMF	How Many Flowers	Student believes the question makes sense to investigate because the question will help determine how many flowers to plant.
HMLB	How Many Ladybugs	Student believes the question makes sense to investigate because the question will help determine how many ladybugs are needed.
LADY	Keep Ladybugs There	Student believes the question makes sense to investigate because the question will help determine how best to keep the ladybugs near broccoli or keep the ladybugs comfortable.
COORD	Coordinated with Additional Question	Student believes the question makes sense to investigate because the question will help provide additional, useful information in conjunction with another question suggested by the student.
WHICH	Which Flower	Student believes the question makes sense to investigate because the question will help determine which flower to plant to attract the ladybugs.
WHERE	Where to Plant	Student believes the question makes sense to investigate because the question will help determine where to plant the flowers.
LOC	Locate Aphids	Student believes the question makes sense to investigate because the question will help determine where the aphids are located.
FULL	Ladybugs Get Full	Student believes the question makes sense to investigate because the question will help determine what it takes for the ladybugs to get full.
TIME	How Much Time	Student believes the question makes sense to investigate because the question will help determine how long something takes.

Not Relevant Cluster

Code	Code Name	Description
NREL	Not Relevant	Student believes that the question does not make sense to investigate because it is not relevant to the idea that the hypothetical student is trying to test.

Non-Epistemic Utility Cluster

Code	Code Name	Description
BEN	Beneficial Outcome	Student believes the question makes sense to investigate because the question will have a generally beneficial outcome.
WORK	Make It Work	Student believes the question makes sense to investigate because the question will help make the experiment happen.
FINDOUT	To Find Out	Student believes the question makes sense to investigate because the question will help the student find out.

Weak Utility Cluster

Code	Code Name	Description
EA	Too Easy	Student believes the question does not make sense to investigate because the question is too easy to investigate.
PRS	Personal Benefit	Student believes the question makes sense to investigate because the question will provide a personal benefit.
UKS	Unknown to Student	Student believes the question makes sense to investigate because the question is unknown to the student.
TTO	To Tell Others	Student believes the question makes sense to investigate because the student can tell other people the answer to the question.
INTS	Interesting	Student believes the question makes sense to investigate because the question is interesting.

Analysis of student rationales by question

In this section I will examine patterns in student rationales by grade cohort for each research question. This will afford a view of the degree to which students consider the epistemic utility of the candidate research questions in their decisions about whether a research question makes sense to investigate. For each research question, I first examine the overall percentage of students in each grade cohort who consider the epistemic utility of the question. This percentage is calculated by summing the percentage of students whose rationale was classified in the “Program Utility” Cluster, the “Specific Epistemic Utility” Cluster, and the “Not Relevant” Cluster. I examine any noteworthy differences in the other cluster percentages between the two grade cohorts and then look at patterns in the individual codes between grade cohorts.

Student rationales for Q1: How many aphids can a ladybug eat in one day?

As discussed in the prior section of this chapter, the first question presented to students (Q1: How many aphids can a ladybug eat in day?) was judged to be worthwhile by a vast majority of students in both grade cohorts. The expert scientist agreed that this would be a worthwhile question to investigate. According to the scientist, the question has particular epistemic utility for two reasons. First, it provides information about the feasibility of using ladybugs as a predator in the fields. If ladybugs do not eat a sufficiently large number of aphids each day, then the required ratio of ladybugs to aphids may be too high for the plan to be feasible. Second, the question can be used to further explore the specific feeding preferences that ladybugs have and to compare different species of ladybugs to find out which species has the largest appetite for aphids. The scientists noted that this experiment would likely be conducted in a laboratory setting and therefore feeding rates may not be identical to those in the field. Therefore, while some students suggest that this question has utility for determining how many ladybugs are needed, the scientists explained that this calculation is not straightforward.

Even though the students largely agreed that this question is worthwhile to investigate, not all students grounded their rationales in the epistemic utility of the research question. A majority of high school students (95%) justified their decision based on the epistemic utility of the question, whereas only 50% of elementary students did so (Table 30).

Table 30: Percentage of Each Rationale Type for Question #1
How many aphids can a ladybug eat in one day?

Rationale Type	**ALL (N=41)**	**ES (N=20)**	**HS (N=21)**
Program Utility	5%	0%	10%
Specific Epistemic Utility	59%	45%	71%
Not Relevant	10%	5%	14%
Non-Epistemic Utility	17%	30%	5%
Weak-Utility	10%	20%	0%
Don't Know	0%	0%	0%
Epistemic Utility (Program+Specific+ Not Relevant)	73%	50%	95%

High school students' rationales for judging this question as worthwhile reflected a range of different kinds of epistemic utility. The two most common forms of utility were "HMF-How Many Flowers" (seven students) and "HMLB-How Many Ladybugs" (four students). For example, one elementary school student explained:

> S (ES_65): *Because if they can eat a certain amount then you know how many ladybugs to put in there like how many flowers to put in there….*

Another student explained why he thought this question (Q1-HowMany) would be worthwhile to investigate:

> S (HS_99): *If you know how many they eat, you know how many ladybugs you should have around that.*

Three high school students suggested an additional question that they thought would be strategic if investigated in conjunction with Q1-HowMany [coded as "COORD-Coordinated with Additional Question"]. These students reasoned that it would also be important to know how many aphids were on the broccoli plants. For example:

> S (HS_100): *First they need to know how many aphids on average are in the whole farm which is probably a lot, but after that if they can know Q1-HowMany, they can know how many ladybugs they should put in the farm….*

Half of the elementary school students' rationales were not based on the epistemic utility of the research question. The most common code among elementary students was "BEN-Beneficial" (six of 20 students). For example, one student explained that:

> S (ES_74): *So she can make the ladybugs eat the aphids so there will be no more aphids.*

Several other elementary students judged the question to be worthwhile because then they could tell other people [TTO]:

> S (ES_75): *So they could let the people know how many aphids that the ladybugs ate.*

These findings suggest that while both groups considered the epistemic utility of the question, they did so to different degrees. The high school students were more likely to consider the epistemic utility of the question than the elementary school students. The epistemic utility suggested by the expert scientist was also accorded more with the utility identified by high school students.

Student rationales for Q2: Does planting flowers next to broccoli cause fewer aphids?

The vast majority of all students thought that Q2 (Does planting flowers next to broccoli cause fewer aphids?) made sense to investigate. This research question most directly tests the hypothetical student's idea that planting flowers will help control the pest problem on the broccoli plants. The expert scientist points out that this research question is correlative in nature. If lower numbers of aphids are observed on broccoli plants that are next to flowers, one cannot necessarily claim that the flowers caused the reduction in the aphid population. The scientist explained:

> *So Q2-NextTo is a correlative experiment. You just see less pests where you have flowers and you say, oh well the flower is where the ladybug lives and the ladybugs kill them. But there are alternate hypotheses so now that we've seen this correlation, we want to backfill this story to understand the ecological mechanism that's actually leading to the lower pest population….*

Thus, the hypothesis that planting flowers next to broccoli plants will result in fewer aphids could be predicated on various alternative theories of the mechanism, (i.e., natural enemy or resource concentration). Even though Q2-NextTo does not specifically mention ladybugs, the natural enemies mechanism is implied in the question and explicitly suggested in the setup of the task. The hypothetical student's understanding is grounded in the natural enemies hypothesis and he or she is therefore trying to engineer a cropping system using floral resource provisioning as a way to attract natural enemies. In this sense, Q2-NextTo most directly tests the students' idea that flowers will result in reduced aphids based on the natural enemies hypothesis.

Even though Q2-NextTo was designed to be the question that most directly tests the hypothetical student's idea, only an analysis of students' decision rationales will provide a sense of whether the students in the study viewed the question as having that epistemic utility. Table 31 shows that there is a wide gap between elementary and high school students as to how they view the epistemic utility of the question. High school students were over five times more likely to view the research question as having epistemic utility than elementary students (67% vs. 13%).

Over half of the high school students (52%) recognized Q2 as having utility for the research program. They either identified the question as the main, overarching question that tests the idea or they viewed the question as important for testing the feasibility of the idea. Four high school students viewed the question in both ways. For example, one student explained that:

> S (HS_103): *That's essentially what she wants to know, she wants to know whether or not the ladybug will actually eat the aphids and protect the broccoli plants and none of the other questions really get to that idea…makes sense to*

> *do it first because these two (Q4-WhichFlowers and Q1-HowMany) are kind of irrelevant if it doesn't work.*

In contrast, only 13% of elementary students attributed their decision to the program utility of the research question.

Table 31: Percentage of Each Rationale Type for Question #2
Does planting flowers next to broccoli cause fewer aphids?

Rationale Type	**ALL (N=37)**	**ES (N=16)**	**HS (N=21)**
Program Utility	35%	13%	52%
Specific Epistemic Utility	8%	0%	14%
Not Relevant	0%	0%	0%
Non-Epistemic Utility	46%	75%	24%
Weak-Utility	11%	13%	10%
Don't Know	0%	0%	0%
Epistemic Utility (Program+Specific+ Not Relevant)	43%	13%	67%

Elementary school students were much more likely to view the Q2-NextTo question as generally beneficial [BEN]. They thought that the research question was a good one to investigate because the outcome of fewer aphids is beneficial. They did not specifically consider the epistemic utility of the question in the same way that the high school students did. For example, one student explained their rationale as follows:

> S (ES_63): *I think I chose that one (Q2-NextTo) because it sounds more better than the other ones and fewer aphids is better than having a lot of them so yeah.*

One possible explanation of the difference in attribution of epistemic utility between high school and elementary school students is the difference in their recognition of the relationship between the question (Q2-NextTo) and the idea being tested. It is possible that the students did not appreciate the epistemic power of Q2-NextTo because the main mechanism (ladybugs) is not explicitly part of the question. Another possible explanation is that younger students do not distinguish Q2-NextTo from the act of planting flowers. Because "planting flowers" is part of the question, students view the question as a logistical step in the process of attracting ladybugs, rather than as a direct correlative experiment. I will return to this conjecture and expand in more detail later in this chapter.

Student rationales for Q3: How many spots do ladybugs have on their backs?

A small minority of elementary school students and no high school students thought this question was worthwhile to investigate. The scientists agreed that this question would not make sense to investigate. The question focuses on a superficial feature of the ladybugs and does not have epistemic utility in the context of this research project. The scientists noted that the number of spots on a ladybug's back varies by species but there are also many species that have the same number of spots, so it is not a differentiator.

Table 32: Percentage of Each Rationale Type for Question #3
How many spots do ladybugs have on their backs?

Rationale Type	**ALL (N=42)**	**ES (N=21)**	**HS (N=21)**
Program Utility	0%	0%	0%
Specific Epistemic Utility	0%	0%	0%
Not Relevant	88%	76%	100%
Non-Epistemic Utility	0%	0%	0%
Weak-Utility	12%	24%	0%
Don't Know	0%	0%	0%
Epistemic Utility (Program+Specific+ Not Relevant)	88%	76%	100%

The high school students were unanimous in their rationales for their Q3 decision (see Table 32). They all explained their decision on the basis that the question was not relevant to testing the student's idea. The following is a typical response by a high school student:

> S (HS_103): *It doesn't relate to what she wants to know…it doesn't relate to like ladybugs eating aphids, it talks about their appearance and that's not really what the question was.*

While 76% of elementary students also considered the question to be irrelevant, there were five students who gave an alternative rationale for their decisions. Two students thought that the question did make sense to investigate because they thought the question was interesting:

> S (ES_76): *Because he must be curious about the ladybug's spots.*

Another two elementary students rejected the question but not because it was irrelevant. Rather they thought the question was too easy:

> S (ES_71): *He could research about it, but it's a very easy one. To be easy, it's very easy. And if you want it to be hard, it's not easy, so hard. So that's why every time I want it to be hard.*

On the whole, the students in this study identified Q3-Spots as a question that did not make sense to investigate. While related to ladybugs, the particular focus on ladybugs' spots was largely deemed irrelevant to the issue of pests and broccoli. While they were not unanimous, it is noteworthy that the vast majority of elementary students were not swayed by the superficial connection to ladybugs.

Student rationales for Q4: Which flowers do ladybugs like best?

Of all the candidate research questions, Q4 (Which flowers do ladybugs like best?) had the highest selection rate by both grade cohorts. The expert scientist agreed that this question would make sense to investigate. According to the expert scientist, this question has specific epistemic utility because certain natural enemies are attracted to specific types of flowers. Investigating this question is useful for determining which flowers to plant to attract the natural enemy (in this case, ladybugs.) The expert scientist suggested using choice tests in a lab to investigate this question as a complement to field observations.

Both grade cohorts of students also attributed epistemic utility to this research question (Table 33). 89% of elementary school students and 95% of high school students based their rationale on the epistemic utility of the question. Nearly all of the students (32 of 39) explained that the question would inform the decision of which flower to plant to attract the ladybugs. For example, one typical student explained:

> S (HS_92): *You should know that so you can know which flowers to put next to the broccoli.*

The second most common rationale "LADY-Keep Ladybugs There" was very similar. These students explained that the proper flower would help keep the ladybugs there by providing a suitable habitat.

Table 33: Percentage of Each Rationale Type for Question #4
Which flowers do ladybugs like best?

Rationale Type	ALL (N=39)	ES (N=18)	HS (N=21)
Program Utility	0%	0%	0%
Specific Epistemic Utility	92%	89%	95%
Not Relevant	0%	0%	0%
Non-Epistemic Utility	5%	6%	5%
Weak-Utility	0%	0%	0%
Don't Know	3%	6%	0%
Epistemic Utility (Program+Specific+ Not Relevant)	92%	89%	95%

While there was widespread consensus about the epistemic utility of this question, there were several students who initially did not view the research question as worthwhile. For example, note the following exchange between student and interviewer:

> I: *Why don't you think Q4-WhichFlowers would be a good question?*
> S (HS_104): *Cause I think this [points to flower icon] is the best one right?*
> I: *I don't know, maybe.*
> S (HS_104): *Okay, I thought this was the best one.*
> I: *That's just an example of a flower.*
> S (HS_104): *Okay, oh yeah you can find…this would be important if you didn't know which one is the best.*

The student interpreted the flower icon as *the* flower that the ladybug prefers. Since the student viewed the flower type as already known, he did not consider the question worthwhile to investigate. In the students' mind, the knowledge of which flower the ladybug likes was certain (it was the particular flower depicted in the icon). However, the interviewer explained that the icon flower was just a placeholder example and was not meant to be a specific flower that the ladybug likes. The student immediately changed his mind, acknowledging the importance of knowing which flower the ladybug likes best. This example highlights the importance of examining the issue of uncertainty in the minds of the students as they evaluate epistemic utility. One mediating factor in their judgment is what they already know to be true, or what they think they know to be true. Prior knowledge is a crucial determinant of one's assessment of epistemic utility.

Student rationales for Q5: Which part of broccoli plants do aphids eat first?

The final question in the Ladybug Task was Q5 (Which part of broccoli plants do aphids eat first?). The expert scientist thought this question was interesting but could not identify any particular epistemic utility. Among students, this question had the most varying responses of all of the research questions. Half of the students thought the research question made sense to investigate. However, there was a wide grade cohort gap in the percentage of students who selected the question (ES =29%; HS=74%). By examining the rationales that students gave for their decision, we can begin to explain why the two grade cohorts differed.

Table 34: Percentages of Each Rationale Type for Question #5
Which part of the broccoli plant do ladybugs eat first?

Rationale Type	**ALL (N=40)**	**ES (N=19)**	**HS (N=20)**
Program Utility	0%	0%	0%
Specific Epistemic Utility	35%	47%	24%
Not Relevant	45%	16%	71%
Non-Epistemic Utility	3%	5%	0%
Weak-Utility	10%	16%	5%
Don't Know	8%	16%	0%
Epistemic Utility (Program+Specific+ Not Relevant)	80%	63%	95%

In Table 34, elementary students attributed specific epistemic utility to the question at twice the rate of high school students (47% vs. 24%). The most common rationale given by elementary students was “WHERE-Where to Plant.” Here is an example of what one student said:

> S (ES_74): *So she could put it maybe close to the flowers, near the flowers, so ladybugs just eat the aphids.*

This student, along with five others, interpreted the research question as a question about the flower’s physical proximity to the broccoli plant. They interpreted the phrase “next to” as referring to physical proximity. Therefore they viewed this question as having epistemic utility for determining how close to the broccoli they would plant the flower. In this sense, the question has potentially important epistemic utility.

In contrast, the majority (71%) of high school students thought this question was not relevant to the floral resource provisioning idea. One student explained her reasoning:

> S (HS_100): *No matter what part they eat first, I mean, they're there, you know? And the ladybugs can eat them so, it doesn't really matter what part they get to first, just as long as, you know, you can get the ladybugs near them.*

Summary of student rationales across research questions

Even though students may agree that a question makes sense to investigate, they might make their decision for very different reasons. We have seen evidence that high school students were more likely to base their decisions on the epistemic utility of the question. While elementary students also identified the epistemic utility of questions, they were less likely to do so than high school students. The elementary school students had reasons for deciding if a question made sense to investigate, but were less likely to pinpoint a specific epistemic benefit from investigating the question. However, the results from this analysis also indicate that most students in the study, including elementary school students, were able to identify and discard a question that was only superficially related to the situation.

How Do Students Sequence Research Questions to Investigate an Idea?

In addition to explaining their rationales for their decisions about which questions made sense to investigate, students were also asked to sequence the worthwhile questions in the order they thought the questions should be investigated. In this section, I examine patterns in the ways that students sequenced the questions.

In typical science classrooms, students use the "scientific method" to investigate a single question. The scientific method is portrayed as a linear, step-by-step process driven by a specific question and resulting in a single answer. However, in professional science, scientists coordinate multiple research questions to build an understanding of the phenomena under study. One research question may be more fundamental, but oftentimes multiple research questions (or sub-questions) are also required to investigate the larger question. A set of research questions can make individual contributions that, taken together, constitute a research program.

The expert scientist interviewed for this study provided his view of a strategic sequence and rationale for investigating these questions. He viewed Q2-NextTo as the main, overarching question because it most directly addresses the idea that the hypothetical student has for reducing the pest population in the broccoli crop. However, the scientist thought it made the most sense to investigate the flower preference (Q4-WhichFlowers) and ladybug appetite (Q1-HowMany) before setting up a field experiment (Q2-NextTo). In his words:

> *Based on information that you derive from the first one (Q1-HowMany) you have been able to establish which species of ladybug is going to have the most impact on these pests or seems to prefer it in the lab, and you also have an idea of the flowers that the ladybug is attracted to (Q4-WhichFlowers) so you have a lot better chance of setting up an experiment that is going to do something.*

This logic is based on the epistemic utility of each research question. Before investigating the main question (Q2-NextTo) it is strategic to find out other information first. In this sense, Q1-HowMany and Q4-WhichFlowers have particular epistemic utility as well as epistemic contingency in relation to Q2-NextTo. Therefore, particular attention is paid towards how students positioned Q2-NextTo relative to the other questions in the sequence and the rationales they provided for their question sequence.

Among students' sequences of the research questions, there were four permutations that varied based on the relative position of Q2-NextTo (see Figure 14). In the first permutation, Q2-NextTo was the first question investigated and the other questions followed in various orders. In the second permutation, Q2-NextTo was the last question investigated. In the third permutation, other questions were asked before and after Q2-NextTo. In the fourth permutation, Q2-NextTo was not selected as part of the sequence.

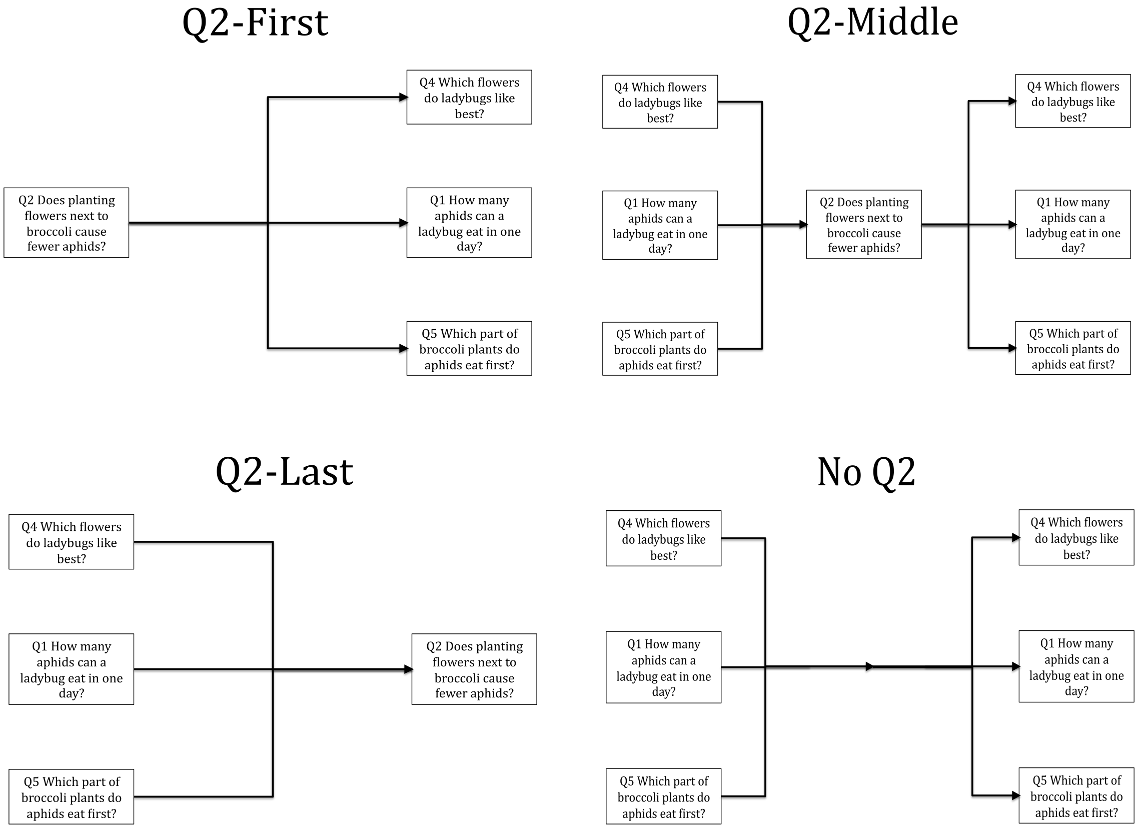


Figure 14: Question Sequence Permutations (Hinging on Q2-NextTo)

In this study, all high school students (N=21) were asked to sequence the research questions that they deemed worthwhile. However, only half of the elementary school students (N=12) were asked to do the same. The Ladybug Task was challenging for the younger students. In cases where the student seemed taxed, the interviewer skipped the sequencing portion of the task. Therefore, only partial results are available for the elementary school cohort. Nonetheless, the inclusion of the analysis of elementary school students is still useful for comparison purposes.

Table 35 shows the range of question sequence patterns expressed by elementary and high school students. The frequency of each pattern ("Count") and proportion of students in each sequence permutation category ("Category Count") is shown. A total of 20 unique sequences were observed across the 33 student responses. The most common permutation was Q2-NextTo, Q4-WhichFlowers, Q1-HowMany. There were three other sequences that were shared by three students each. Two sequences were shared by two students each. Only one student expressed each of the remaining sequences. This distribution reflects a remarkable diversity in question sequences without much clustering (besides Q2-NextTo, Q4-WhichFlowers, Q1-HowMany).

Table 35: Frequency of Question Sequence Permutation by Grade Cohort

Category	Pattern	Count (ES)	Count (HS)	Total Count	Category Count (ES)	Category Count (HS)	Category Count (ALL)
Q2-Last	Q1, Q4, Q5, Q2		1	2	2 (17%)	8 (38%)	10 (33%)
	Q1, Q5, Q4, Q2	1		1			
	Q4, Q1, Q2		3	3			
	Q4, Q1, Q5, Q2	1	1	2			
	Q4, Q2		1	1			
	Q5, Q1, Q4, Q2		1	1			
Q2-First	Q2, Q1, Q4		1	1	3 (25%)	9 (43%)	12 (36%)
	Q2, Q1, Q5, Q3	1		1			
	Q2, Q4		1	1			
	Q2, Q4, Q1		6	6			
	Q2, Q4, Q5, Q1	1		1			
	Q2, Q5, Q4		1	1			
	Q2, Q5, Q4, Q1	1		1			
Q2-Middle	Q1, Q2, Q4		1	1	4 (33%)	2 (10%)	6 (18%)
	Q4, Q2, Q1	2	1	3			
	Q4, Q5, Q2, Q1	1		1			
	Q5, Q2, Q1, Q4	1		1			
No Q2	Q4, Q1		1	1	3 (25%)	2 (10%)	5 (15%)
	Q4, Q1, Q5	2	1	3			
	Q5, Q4, Q1	1		1			
		N=12	N=21	N=33	N=12	N=21	N=33

With respect to the permutations of sequencing Q2-NextTo, high school students most commonly placed Q2-NextTo either first (43%) or last (38%). Only two high school students placed Q2-NextTo in the middle and another two high school students did not select Q2-NextTo as worthwhile. By contrast, over half of the elementary school students placed Q2-NextTo in the middle or not at all. This finding represents a striking difference between the elementary and high school students. High school students were almost twice as likely as elementary school students to place Q2-NextTo either first or last (81% for high school and 42% for elementary school students.) This finding reflects the earlier finding about the epistemic utility attributed to Q2-NextTo by high school students compared to elementary school students. In particular, high school students were more likely to attribute “program utility” to Q2-NextTo, thus informing how they would sequence the research questions. But to explain more adequately the sequencing patterns, I turn to an analysis of students’ rationales for why they sequenced the questions in the way that they did.

On What Basis Do Students Rationalize the Coordination of Research Questions?

In this section I examine students’ rationales for why they sequenced the research questions in the way that they did. This analysis offers a view of how students do or do not consider the epistemic utility of questions when thinking about how the questions may relate to each other. In a coordinated research program, scientists often investigate multiple research questions that have epistemic ramifications for each other. Knowledge gained from one investigation can lead to another investigation.

Schema for sequence rationales types

In the Ladybug Task, students were asked to sequence questions they thought made sense to investigate in the order they thought the questions should be investigated. They were also asked to explain why they sequenced the questions in that particular order. By focusing on the specific rationales students gave for the sequence placement of each question, I developed a coding schema for categorizing the different types of sequence rationales. As the schema emerged, I applied it to the sequence rationales and revised as needed until it fit with the student responses.

Table 36 shows the coding schema of the different types of rationales students gave for their question sequences. On a basic level, some students sequenced the research questions based on the order of importance of the questions rather than on specific contingency between the questions. For example, one student explained why he would investigate Q1-HowMany last:

> S (ES_80): *He could check that one [Q1-HowMany] last because these [the other questions] are the most important ones that he needs to find out.*

The student is thinking about how important the questions are relative to each other and then sequencing them accordingly from most important to least.

Table 36: Question Sequence Rationale Types

Sequence Rationale Type	**Description**
Importance	Sequence rationale based on the importance of the research question. The questions are prioritized in order of importance.
Logistical Sequence	Sequence rationale based on a step-by-step logistical process.
Test Then Proceed	Sequence rationale based on prioritizing specific question(s) that can determine if idea is going to work before proceeding with remaining questions.
Summing Up / Building Towards	Sequence rationale based on prioritizing questions that provide requisite information for investigating subsequent questions.
No contingency	Sequence rationale is not based on any particular contingency between questions. Rationale provided insufficiently explains sequence.
No opportunity	Students were not asked to explain sequence.

Other students sequenced the research questions based on some degree of contingency or dependency between one question and the next. For some students, this contingency was conceived primarily in terms of steps of a process. For example, the following student explains her rationale in logistical terms:

> I: *Why should that be the order?*
> S (ES_90): *Because we need to find out what kind of plant does the ladybug like and then we find the plant then we plant it around the broccoli, this one (points to Q2-NextTo), and then we find out how many aphids a ladybug can eat in one day.*

The student interprets Q2-NextTo as the step when the flowers get planted. The student reasons that only once the flowers are planted and the ladybugs arrive can one find out how many aphids a ladybug can eat in one day (Q1-HowMany). The student is thinking in logistical terms, as in which question needs to be investigated in order to have the (practical) opportunity to investigate the next question. While contingency plays a role in the sequence, the student is not thinking in terms of epistemic utility.

In contrast, the epistemic utility of the research questions does inform how other students rationalize their sequence. The knowledge gained from investigating one question can have implications for the investigation of the other questions. This epistemic contingency was manifested in two primary ways. In some cases students explained their rationale for investigating certain questions first because they thought it was important to know the answer to the question before proceeding with the remaining questions. For example, this student said that it was important to first test a particular question to see if it works:

> S (HS_97): *First of all I choose Q2-NextTo, because if they don't cause fewer aphids, then what's the point of doing this whole thing? So, why go through all this work before doing that?*

This student thinks that Q2-NextTo is important to investigate first because if the idea doesn't work, then the other questions do not make sense to investigate. Students like this one wanted to test the feasibility of the idea before proceeding with the other questions.

The second form of epistemic contingency was characterized by sequence rationales oriented around "summing up" or "building towards" a larger question. These students sequenced the research questions such that the questions investigated first would culminate or build towards the final research question. The students viewed the initial questions as requisite components of the larger, final question. For example:

> I: *What about Q2-NextTo? [Why would you investigate that question last?]*
> S (HS_111): *Because I think that all of these [other questions] are leading up to the experiment and then this [Q2-NextTo] would be the experiment itself. So you want to get all your information in your hypothesis before you do the experiment and I think this [Q2-NextTo] would be your conclusion almost.*
> I: *Okay, so once you've done these three questions [Q1-HowMany, Q4-WhichFlowers, Q5-EatFirst] do you still need to do something for this one [Q2-NextTo] or do you kind of automatically know it?*
> S (HS_111): *I think you would have to do the experiment by planting everything.*

This student views Q2-NextTo as the main question and the other questions contribute to the Q2-NextTo investigation. The student still believes it is important to actually investigate Q2-NextTo, once the other questions had been investigated first. Other students coded as "Summing Up / Building Towards" also agreed that Q2-NextTo would need to be investigated.

In a handful of cases, students' rationales referenced more than one form of contingency. In these cases, preference was given to "Building Towards" and "Test Then Proceed" over other rationales since these categories reflect higher sophistication. And finally, there were students who did not offer a rationale that explained their question sequence or were not asked to explain their rationale.

Patterns in sequence rationales by grade cohort

Elementary school and high school students use different rationales to explain their research question sequences (see Table 37). Elementary school students more commonly offer logistical explanations for their question sequences. In contrast, over half of the high school students explained their rationales with consideration of epistemic contingency ("Test Then Proceed" or "Summing Up / Building Towards"). These results indicate that elementary students were more likely to view the question sequences as steps of process rather than as a set of coordinated opportunities to build useful knowledge about the phenomenon.

Table 37: Prevalence of Question Sequence Rationale Type by Grade Cohort

Sequence Rationale Type	Elementary School Students	High School Students
Importance	8%	5%
Logistical Sequence	67%	19%
Test Then Proceed	0%	38%
Summing Up / Building Towards	0%	24%
No contingency	17%	14%
No opportunity	8%	0%
	N=12	N=21

Given that Q2-NextTo has a distinct epistemic role in this task, it is worth exploring the relationship between how students placed Q2-NextTo in the sequence and their corresponding sequence rationale. To investigate these patterns, Table 38 shows the frequency counts of question sequence permutations by question sequence rationale by grade cohort. This analysis builds on the analysis in the prior section regarding the different sequence permutations that hinged around the placement of Q2-NextTo.

There are several important patterns to notice in this analysis. First, both grade cohorts used the "Logistical" rationale to explain a variety of different sequence permutations. The "Logistical" rationale is not associated with any particular sequence of research questions. This finding suggests Q2-NextTo does not take on a consistent significance across the students when they are thinking about the questions in terms of logistical sequence.

The second noteworthy pattern is that for high school students there was a relationship between the sequence permutation (Q2-First & Q2-Last) and the sequence rationale type. All high school students who explained their sequence using the "Test Then Proceed" rationale thought that Q2 should be investigated first. All students who explained their sequence with the "Summing Up / Build Towards" rationale thought that Q2-NextTo should be investigated last.

Table 38: Counts of Sequence Permutations by Rationale

		Importance	Logistical Sequence	Test Then Proceed	Summing Up / Build Towards	No Contingency	No Opportunity	Total
ES	Q2-First	1	1			1		3
	Q2-Middle		3				1	4
	Q2-Last		2					2
	No Q2		2			1		3
Total ES		1	8	0	0	2	1	12
HS	Q2-First	1		8				9
	Q2-Middle		1			1		2
	Q2-Last		2		5	1		8
	No Q2		1			1		2
Total HS		1	4	8	5	3	0	21

These findings suggest that there is a close relationship between how students interpret the questions and how they think the questions fit together. Students who view Q2-NextTo as the main question either think it is important to investigate first or last. Those who select Q2-NextTo first explain their sequencing in terms of testing the main idea, to see if the idea will work before proceeding. Those who select Q2-NextTo last think it is important to investigate the other sub-questions first, to inform the investigation of the main question.

However, real world science does not necessarily always follow the most logical pathway. There may be a wide range of circumstantial conditions that dictate how a research program proceeds. In fact, the expert scientist interviewed for this research study explained that his team's research program began with Q2-NextTo. The expert scientist, who studies floral resource provisioning with wasps in vineyards, started with Q2-NextTo largely due to the critical importance of collaborating with the vineyard growers:

> *We began with this last step (Q2-NextTo) where our work is in vineyards and the growers that approached us were interested in experimenting with habitat diversification and so we first said, before we embark on this huge study and put all these grad students on it, we just did some field evaluations, like vineyards that had flowers planted in them, compare that to a control plant*

> *with no flowers and then we started organizing some growers to plant the same set of flowers and I guess also the collaborative nature of this, we really want it to be something that they would adopt and so before going into a lab and working on what flowers the bugs like we started with the flowers that the growers like in the sense that they are actually able to manage them effectively in a vineyard in the northern part of California. So they needed stuff that was sown in the fall, needed stuff that was drought-tolerant, that would bloom one after another in a vineyard.* [Expert Scientist]

The expert scientist first conducted field experiments similar to Q2-NextTo and plans to subsequently conduct laboratory experiments around floral preference and predator appetite to further examine the mechanism for the correlative findings. In this case, the sequence of research investigations conducted by the expert scientist matched those students who suggested that Q2-NextTo be investigated first to test feasibility. Of course the expert scientists did already have a high degree of background knowledge about predatory relationships and could consult published resources about floral preferences.

Chapter Discussion

In this chapter, I analyzed student rationales about which questions made sense to investigate to test an idea in an ecological context. Of particular interest was the extent to which students considered the epistemic utility of the research questions. Findings from this analysis suggest that high school students more commonly considered the epistemic utility of the research question in their rationales. Understandably, students' evaluations about the value of investigating a particular question depended on their interpretation of what the research question was about. In the case of Q2-NextTo, elementary school students more commonly interpreted the question in terms of the physical proximity of the flower to the broccoli plants. This interpretation was a direct determinant of the epistemic utility that they voiced in their rationale. High school students were more likely to identify Q2-NextTo as the main question that most directly tests the floral resource provisioning idea. For students who did not make this connection, the epistemic utility of the question may be less transparent.

Another factor that played a role in students' decisions and rationales was students' current knowledge base. In the analysis for Q4-WhichFlowers, the attribution of epistemic utility was in part a function of what information the student considered to be already known or certain. If the ladybug's flower preferences are already known (as indicated by the icon flower) then investigating that question does not have epistemic utility.

This task was designed to build students' content knowledge about the ecological principles and specific organisms in this example. The task explicitly presented the hypothesis of natural enemies as the idea being tested. The interviewer explicitly reminded the student that the idea is uncertain, the hypothetical student "doesn't know if this idea will work." However, as evidenced by the flower choice, students still considered different aspects of the scenario as known or unknown, and their views made a difference for their assessments of the utility of particular research questions.

In the Ladybug Task, students also considered how to coordinate multiple research questions into a research program. This prompt invited students to think about the relative utility of each question and to sequence them in strategic ways. A major affordance of this approach is that it highlights how knowledge can be constructed through the pursuit of coordinated questions. In science, questions can have contingency, such that finding out the answer to one question is important to pursuing the next one. In the analysis of student thinking from the Ladybug Task, there is evidence that students also reasoned about how questions are coordinated and in some cases the questions are contingent on each other. However, the high school students more commonly recognized the primary question (Q2-NextTo) as having central importance to the research agenda and thus placed the question strategically in the research question sequence.

Chapter 6: Conclusion

This research introduces and elaborates the idea of "worthwhileness" as a strategic construct in the field of epistemic cognition, worthy of empirical research. The notion of worthwhileness is a crucial aspect of epistemic reasoning. Thinking about what makes a question worthwhile is fundamentally epistemological in nature. Scientific questions are an epistemic form, and epistemic criteria used in evaluating research questions are an implicit part of one's practical epistemology of science.

Determining whether a research question is worthwhile is also reflected in authentic scientific practice in professional science. Scientists develop research questions to advance theory, and worthwhile questions that have epistemic utility are fundamental to that process. In practice, scientists must also make a case for why their research is worthwhile in order to acquire funding to support their work. Similarly, scientific journals require that article submissions explain why the research is worthwhile and what contribution it makes to the field.

The study of students' beliefs about worthwhile questions is highly strategic. By examining students' conceptualizations of what makes a research question worthwhile, we can better understand their epistemic criteria, which is a reflection of students' personal epistemologies of science. Evaluating a research question is an opportunity to reflect on the goals of science and the nature of scientific contributions. It is also possible that a strategic focus on the epistemic contribution of a research question could lead to more powerful student-generated questions in school science.

Investigating worthwhileness

To examine the construct of worthwhileness, this dissertation study used novel research tasks for engaging students in epistemic games that activated their epistemic criteria. Historically, many researchers have used traditional instruments such as fixed choice questionnaires and surveys to measure students' attitudes and beliefs about the nature of science (Lederman, Wade, & Bell, 1998) while others have used interview protocols (e.g., Carey et al., 1989). These instruments have typically focused on students' formal epistemologies (i.e., beliefs about professional science.) As a complement to research on students' formal epistemologies, Sandoval (2005) has argued for the importance of research that examines students' tacit epistemological beliefs as reflected in their engagement in scientific practice which he refers to as students' *practical* epistemologies of science.

One way that students' practical epistemologies are manifested is in the implicit criteria that students use to evaluate particular epistemic forms in the context of scientific practice. Therefore, the tasks in this study provide students with opportunities to evaluate and critique examples of scientific research questions as they make decisions about which questions are worthwhile to investigate. For example, the Funding Allocation task asks students to evaluate a set of scientific research questions and to decide which questions warrant funding.

The Funding Allocation Task extends the work of Hogan & Maglienti (2001) who investigated epistemic criteria about scientific conclusions. In the Hogan & Maglienti study, students (and adults) were asked to rate the validity of a set of hypothetical conclusions derived from data. The implicit task demands were for students to articulate the criteria they used in rating the scientific conclusions. The Funding Allocation Task also elicits students' epistemic criteria as reflected in students' evaluations, but the Funding Task targets criteria about scientific research questions. The funding context also provides a purposeful reason to motivate students' decisions.

Another affordance of the Funding Allocation Task is that it invites students to both reflect on how they made their funding decisions and to articulate the criteria that they used. Pluta et al (2011) used a similar tactic to investigate students' epistemic criteria for good scientific models. Pluta et al. gave students a packet full of different examples and non-examples of scientific models as well as contrasting models of specific phenomena (i.e., global warming, plant growth, etc.) Students were asked to compare and contrast the different examples and to discuss with their classmates what makes one model better than another. The authors then collected and analyzed a list of criteria for good models from each student. This dissertation study applies a similar approach to investigating students' epistemic criteria for scientific research questions. However, the Funding Task also affords a comparison between the implicit criteria reflected in students' rationales for each funding decision and the explicit articulation of criteria students express retrospectively. In this way, the Funding Task extended the research Pluta et al. conducted on epistemic criteria for scientific models.

The Driver et al. (1996) study on students' epistemologies of science included one task that most closely resembles the tasks used in this study. In the Driver et al. study, students were given a set of research questions and asked which ones were "scientific." The interviewers told students that "scientific" questions were questions that scientists might want to find out more about. Driver et al. analyzed the bases of students' judgments to determine how students viewed the purpose of scientific work.

While the Funding Allocation Task also asks students to make judgments about a set of research questions, the task is different from the one used by Driver et al. in several important ways. First, the task demands of the Funding Allocation Task are different. Rather than deciding whether a question is "scientific," the students are asked to decide which questions they think are worthwhile to investigate. This task elicits the full scope of students' criteria for worthwhile questions. While some students explained their decisions on the basis that the question was "scientific," that was only one of many criteria reflected in students' rationales.

A second important distinction between Driver's tasks and the Funding Allocation Task in this dissertation is that the Funding Allocation Task explicitly asked the students to explain why *they* (the students) thought that a scientific research question was worthwhile to investigate. This distinction is subtle yet crucial. Although the Funding Allocation Task involved the funding of scientific research questions for investigation by professional scientists, the task demands positioned students to make decisions based on their own criteria for what makes a question worthwhile to investigate. Therefore, in contrast to the Driver et al. task, student responses to the Funding Allocation Task reflect students' own practical epistemologies of science.

This study also includes a task designed to explicitly distinguish between the criteria that students apply in a professional science context from the criteria they believe should be used in a school science context. The Science Fair Task afforded a comparison between students' conceptions of worthwhile scientific research questions for professional scientists and for school science. Students suggested criteria for use by judges in awarding a prize for the best scientific research question at a school science fair. Differences in the weighting of criteria students used in that context compared to the criteria they listed for making funding decisions reflects differences in their images of professional versus school science, a distinction identified by Sandoval (2005).

The Ladybug Task represents a departure from prior work in epistemic cognition. As a complement to the other tasks in the interview, the Ladybug Task is specifically designed to examine how students reason about the epistemic utility of a scientific research question within the context of a particular research program or problem. This study conceptualizes the epistemic utility of a question as the contribution that investigating the question would make towards the construction of scientific knowledge about the topic or phenomena under study. The epistemic utility of a research question could be one of several epistemic criteria for what makes a scientific research question worthwhile to investigate.

To position students to consider the epistemic utility of a scientific research question, the Ladybug Task situates students in a content-rich, goal-oriented context that invites students to coordinate multiple research questions to investigate an idea. All of the research questions target the same scientific domain and can therefore be compared and contrasted with each other to determine their relative epistemic utility. This type of epistemic reasoning was also evident when students explained how they thought the questions should be strategically sequenced to investigate the idea under study.

While the Funding Task afforded a view of students' general epistemic criteria and funding priorities, the Ladybug Task was particularly effective at eliciting students' ideas about the epistemic utility of scientific research questions. The specific domain context allowed for a finer-grain analysis of relative epistemic utility between questions. In addition, the coordination of multiple questions was fruitful for understanding how students considered the epistemic contingency between questions.

The novel tasks used in this dissertation explore students' epistemic criteria about worthwhile scientific questions and extend prior research on students' criteria for other epistemic forms. The dissertation tasks were specifically designed to elicit students' epistemic criteria about worthwhile scientific research questions in a variety of different contexts. While the particular contexts explored in this study do not reflect all possible contexts in which students might apply epistemic criteria, the specific tasks and contexts do expose how students apply criteria differently depending on the context.

Empirical findings

This research study yielded several noteworthy findings. Generally, findings from this study suggest that students had ideas about what makes a research question worthwhile and were able to engage in the epistemic game of evaluating research questions. High school students were more readily able to engage in this epistemic game, and successful engagement with the tasks was more variable for elementary school students.

In playing the epistemic game, students considered a variety of epistemic criteria including interest, difficulty, knowledge status and the extent to which the question addressed a perceived problem in the world. In both grade cohorts, students' reasoning most heavily weighted consideration of the question's contribution. However, there were significant differences in weighting between the cohorts, with elementary students more commonly considering interest and difficulty criteria compared to high school students.

Another finding from this study was that the explicit and tacit epistemic criteria that students use are not necessarily one and the same. Analysis of students' retrospective criteria suggests that there is variation in what students expressed when evaluating the questions compared to their explicit articulation of those criteria afterwards. Grade level differences indicate that while both cohorts matched to some degree, the high school students' explicit, retrospective criteria most closely reflected the reasoning embedded in their specific funding decisions.

The differences in students' specific funding decisions compared to their articulation of generalized criteria fits with the Ericsson & Simon (1980) research on verbal protocol production. Ericsson & Simon make the case that there is not necessarily a tight match between how students solve problems and how they report that they solved the problems. Students' retrospective description of how they solved a series of problems may be even further removed from the thinking they actually employed. These distinctions are important in the context of epistemic criteria. It is a different cognitive task to explain why you made a decision than to generalize and articulate the criteria that you used. Therefore, tasks that require students to explain their decisions as well as enumerate specific criteria are complementary methods for getting at students' implicit and explicit epistemic beliefs.

Empirical findings from this research suggest that the context in which students apply epistemic criteria can impact which criteria students use. Students often conceptualize school science and professional science quite differently (Driver et al., 1996; Metz, 2008; Sandoval, 2005). In this study, students more frequently reported consideration of the contribution of a research question when they reasoned about worthwhile scientific research questions for professional scientists, compared to reasoning about worthwhile questions in a school science context. This evidence suggests that students' epistemologies of science may take different forms depending on the context, and that students regard school science as a different practice than professional science.

Close analysis of students' epistemic reasoning as they engaged in the study tasks yielded a more refined view of how students conceptualized the nature of a research question's contribution. At some point during the interviews, all students reasoned about a question's contribution in terms of the goals of science including consideration of a question's epistemic and functional utility. While this type of reasoning maps to the Stokes (1997) framework for the goals of science, students rarely considered both criteria (epistemic and functional utility) in the evaluation of a particular question. This finding is reasonable to expect given that, according to Stokes, a false binary between a "quest for understanding" and "considerations of use" still exists as separate goals for science and technology in the professional world.

There is also evidence to suggest that students' evaluations of research questions were mediated by students' content knowledge about the subject as well as their own personal values or world-view. In deciding if a question was worthwhile to investigate, many students compared the target knowledge of the question to what they already knew about the topic or what they perceived was already known about the topic. There were also many cases in which the student rejected a question because the question was based on assumptions that the student knew to be false, reflecting the students' own internal epistemic analysis.

This research also suggests that students can coordinate multiple research questions to create a plan for investigating specific ideas. Many students in the study considered the epistemic utility of a related set of questions. While there was variation among student participants, many students also sequenced the questions on the basis of epistemic contingency. Specifically, some students prioritized testing feasibility by investigating the main question first, while others thought it made more sense to build towards the main question after investigating subordinate ones first. In both cases, students considered the particular epistemic utility of the research questions and thought about contingency between questions when explaining their sequences.

Taken as a whole, the empirical findings from this study suggest that students had ideas about what makes a scientific question worthwhile and that they applied epistemic criteria in their evaluations of worthwhile questions. However, consideration of the questions' contribution varied across students and contexts. When students were asked to critique a set of questions about a particular topic in relation to a specific goal, students' considerations of the epistemic utility of a research question played a larger role in their criteria.

Extending the research literature

This research study makes several contributions to the research literature on epistemic cognition. The construct of *worthwhileness* is itself a contribution to the field of epistemic cognition. The construct offers a new frame for the study of epistemic criteria in its application of the criteria to a particular judgment about the merit of a scientific research question. Prior research has examined students' epistemic criteria in relation to other epistemic forms such as scientific models (Pluta et al., 2011) and scientific conclusions (Hogan & Maglienti, 2001). However, very limited attention has been paid to epistemic criteria for scientific research questions (Driver et al., 1996).

Carey et al. (1989) examined students' ideas about the source of scientists' ideas and their motivations for conducting experiments but the authors did not elicit students' specific criteria about research questions. Similar to the present study, the Driver et al. (1996) study also asked students to evaluate a set of scientific research questions. However, Driver and her colleagues asked students to decide if the questions were "scientific" or not, which the interviewers defined as a question that a scientist would investigate. Students were asked to consider how a scientist would think about the question, rather than making their own judgment about what they think. This study extends the work by Driver et al. by positioning students to evaluate scientific questions on the basis of what *the student* thinks makes a question worthwhile to investigate. This shift allows for an expanded view of the particular epistemic criteria that students apply.

Several prior studies have advocated for the use of criteria for good questions with students (Chin & Kayalvizhi, 2002; Lucas et al., 2005; White et al., 2009; White et al., 2011). While these studies focused on supporting students in generating scientific research questions and include criteria for good questions, they did not examine the extent to which students adopted these criteria as part of their personal epistemologies of science. To extend that prior work, this research study examined criteria expressed by the students themselves.

This study's focus on students' conceptualizations of the contributions of particular research questions in terms of epistemic and functional utility is an unexplored area in the research literature. Furthermore, the explicit focus on the relative epistemic utility of a set of scientific research questions adds a dimension of epistemic reasoning across multiple questions and represents an important new avenue for future research. While the notion that authentic inquiry is often comprised of many questions is acknowledged in the research literature (White, 2011), this study actually investigates how students reason when faced with a set of questions that may relate in different ways to an epistemic goal as part of a coordinated research program.

Finally, this dissertation study begins to investigate the issue of epistemic value, a construct that Chinn (2011) suggested as a focus of future research on epistemic cognition. In their expanded theoretical framework for epistemic cognition, Chinn et al. (2011) argued that epistemic values are part of one's personal epistemology and that future work should examine how students prioritize various epistemic achievements. Along those lines, this study begins to explore the ways in which students value particular epistemic aims and how they reason about questions of varying epistemic utility. Furthermore, students' own personal values were evident in their applications of epistemic criteria to scientific research questions. While more research is needed to fully explore the dimensions of Chinn et al.'s framework, this study represents one step in that exploration.

Instructional implications

This study could have implications for instructional practice in science education. Students' understanding about what makes a scientific research question worthy to investigate can be developed through instruction. Curricula could be designed to support and facilitate the development of this construct in line with normative views of the nature of science and scientific questions. Science educators and curriculum designers might consider taking up the construct of worthwhileness as an object of thought with students. Given that students do have ideas about what makes a worthwhile scientific question, it could be useful to expose those ideas and make them explicit. Building on students' initial epistemic criteria, science instruction could then scaffold ways of considering the epistemic utility and functional utility of a research question. The interview tasks used in this study could be adapted as classroom based activities because they activate students' epistemic criteria and require students to apply those criteria in engaging contexts (i.e., allocating research funding).

Limitations

While there are many contributions of this research, it is also important to note the limitations of the study. First, this research study is an initial exploration of the worthwhileness construct. Further refinement of the tasks and analytic methods would likely provide an even more sophisticated view of students' conceptions. Second, the study was small in scale with a limited number of participants. Therefore, a larger study would be required to generalize the findings to the student population as a whole.

Third, the tasks used in this study focused exclusively on the critique of scientific research questions, not on the generation of questions. Since a goal of science education is to support students in formulating their own research questions, an expansion of this study would need to include opportunities for students to generate their own questions. Lastly, it is important to note that the study did not have an instructional component beyond students' engagement in the tasks themselves. Therefore the findings do not address the epistemic criteria that students might develop given strategic instruction.

Future research directions

There are several productive directions for building on this dissertation research. One direction could include revising the assessment to more deeply explore the relationship between content knowledge, conceptual understanding and reasoning about epistemic utility. While the Ladybug Task begins to investigate these issues, multiple, isomorphic domain contexts could help determine the extent to which evaluations of epistemic utility are a function of a student's conceptual understanding.

A second research direction could be to design instructional interventions to support students' epistemic reasoning in line with the work of White et al. (2009, 2011). The tasks used in this research could be reformulated as whole class activities that encourage students to consider the construct of worthwhileness as an object of thought. It could be strategic to study how these epistemic games function to develop students' epistemic criteria and to determine the extent to which students' more sophisticated criteria result in an improvement in the quality of student-generated scientific research questions.

The intersection between epistemic cognition and research on science education is ripe for further research. While there is an increasing focus on students' epistemic criteria and the scaffolding thereof, the research in this area is still quite limited and there are many contributions to be made. Developing students' epistemologies of science is a central goal of science education and a worthwhile focus for further research.

References

Alfke, D. (1974). Asking operational questions. *Science and Children, 11*, 18-19.

Allison, A. W., & Shrigley, R. L. (1986). Teaching children to ask operational questions in science. *Science Education, 70*(1), 73-80.

Alters, B. J. (1997). Whose nature of science? *Journal of Research in Science Teaching, 34*(1), 39-55.

Blumenfeld, P. C., Soloway, E., Marx, R. W., Krajcik, J., Guzdial, M., & Palinscar, A. (1991). Motivating project-based learning: Sustaining the doing, supporting the learning. *Educational Psychologist, 26*(3&4), 369-398.

Carey, S., Evans, R., Honda, M., Jay, E., & Unger, C. (1989). 'An experiment is when you try it and see if it works': a study of grade 7 students' understanding of the construction of scientific knowledge. *International Journal of Science Education, 11*(Special Issue), 514-529.

Carey, S., & Smith, C. (1993). On understanding the nature of scientific knowledge. *Educational Psychologist, 28*(3), 235-251.

Chin, C., & Chia, L.-G. (2004). Problem-based learning: Using students' questions to drive knowledge construction. *Science Education, 88*, 707-727.

Chin, C., & Kayalvizhi, G. (2002). Posing problems for open investigations: What questions do pupils ask? *Research in Science & Technological Education, 20*(2), 269-287.

Chin, C., & Kayalvizhi, G. (2005). What do pupils think of open science investigations? A study of Singaporean primary 6 pupils. *Educational Research, 47*(1), 107-126.

Chin, C., & Osborne, J. (2008). Students' questions: A potential resource for teaching and learning science. *Studies of Science Education, 44*(1), 1-39.

Chin, C., & Osborne, J. (2010). Supporting argumentation through students' questions: Case studies in science classrooms. *Journal of the Learning Sciences, 19*, 230-284.

Chinn, C. A., & Malhotra, B. A. (2002). Epistemologically authentic inquiry in schools: A theoretical framework for evaluating inquiry tasks. *Science Education, 86*, 175-218.

Chinn, C. A., Buckland, L. A., & Samarapungavan, A. (2011). Expanding the dimensions of epistemic cognition: Arguments from philosophy and pscyhology. *Educational Psychologist, 46*(3), 141-167.

Chouinard, M. M. (2007). Children's questions: A mechanism for cognitive development. *Monographs for the society for research in child development, 72*, 1-108.

Collins, A., & Ferguson, W. (1993). Epistemic forms and epistemic games: structures and strategies to guide inquiry. *Educational Psychologist, 28*(1), 25-42.

Cuccio-Schirripa, S., & Steiner, H. E. (2000). Enhancement and analysis of science question level for middle school students. *Journal of Research in Science Teaching, 37*(2), 210-224.

Deci, E. L., & Ryan, R. M. (1987). The support of autonomy and the control of behavior. *Journal of Personality and Social Psychology, 53*(6), 1024-1037.

Driver, R., Leach, J., Millar, R., & Scott, P. (1996). *Young people's images of science*. Bristol, PA: Open University Press.

Duschl, R. A. (2007). Quality argumentation and epistemic criteria. *Contemporary Trends and Issues in Science Education, 35*(2), 159-175.

Duschl, R. A., & Osborne, J. (2002). Supporting and promoting argumentation discourse in science education. *Studies of Science Education, 38*, 39-72.

Duschl, R. A., Schweingruber, H. A., & Shouse, A. W. (Eds.). (2007). *Taking science to school: learning and teaching science in grades k-8*: The National Academies Press.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review, 87*(3), 215-251.

Germann, P. J., Haskins, S., & Auls, S. (1996). Analysis of nine high school biology laboratory manuals: Promoting scientific inquiry. *Journal of Research in Science Teaching, 33*(5), 475-499.

Hofer, B. 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.

Hofstein, A., Navon, O., Kipnis, M., & Mamlok-Naaman, R. (2005). Developing students' ability to ask more and better questions resulting from inquiry-type chemistry laboratories. *Journal of Research in Science Teaching, 42*(7), 791-806.

Hogan, K., & Maglienti, M. (2001). Comparing the epistemological underpinnings of students' and scientists' reasoning about conclusions. *Journal of Research in Science Teaching, 38*(6), 663-687.

Katz, I., & Assor, A. (2007). When choice motivates and when it does not. *Educational Psychology Review, 19*(4), 429-442.

Katz, L. G., & Chard, S. C. (1998). Issues in selecting topics for projects. *ERIC Digest, ED424031*.

Keys, C. W. (1998). A study of grade six students generating questions and plans for open-ended science investigations. *Research in Science Education, 28*(3), 301-316.

Khishfe, R., & Abd-El-Khalick, F. (2002). Influence of explicit and reflective versus implicit inquiry-oriented instruction on sixth graders' views of nature of science. *Journal of Research in Science Teaching, 39*(7), 551-578.

Kitchener, K. (1983). Cognition, metacognition, and epistemic cognition. *Human Development, 26*, 222-232.

Kolsto, S. D. (2001). Scientific literacy for citizenship: Tools for dealing with the science dimension of controversial socioscientific issues. *Science Education, 85*, 291-310.

Kruglanski, A. W. (1990). Lay epistemic theory in social-cognitive psychology. *Psychological Inquiry, 1*(3), 181-197.

Lederman, N. G. (2007). Nature of science: past, present, and future. In S. K. Abell & N. G. Lederman (Eds.), *Handbook of research on science education* (pp. 831-880). Mahwah, NJ: Lawrence Erlbaum Associates.

Lederman, N. G., Abd-El-Khalick, F., Bell, R. L., & Schwartz, R. S. (2002). Views of nature of science questionnaire: Toward valid and meaningful assessment of learners' conceptions of nature of science. *Journal of Research in Science Teaching, 39*(6), 497-521.

Lederman, N. G., Wade, P. D., & Bell, R. L. (1998). Assessing the nature of science: What is the nature of our assessments? *Science and Education, 7*, 595-615.

Lehrer, R., & Schauble, L. (2008). Supporting development of the epistemology of inquiry. *Cognitive Development, 21*, 512-529.

Lucas, D., Broderick, N., Lehrer, R., & Bohanan, R. (2005). Making the grounds of scientific inquiry visible in the classroom. *Science Scope, 29*(3), 39-42.

McComas, W. F., & Olson, J. K. (1998). The nature of science in international science education standards documents. In W. F. McComas (Ed.), *The nature of science in science education: Rationales and strategies* (pp. 41-52). Dordrecht: Kluwer.

Metz, K. (2004). Children's understanding of scientific inquiry: Their conceptualization of uncertainty in investigations of their own design. *Cognition and Instruction, 22*(2), 219-290.

Metz, K. E. (2008). Narrowing the gulf between the practices of science and the elementary school science classroom. *Elementary School Journal, 109*(2).

Metz, K. E. (2011). Disentangling robust developmental constraints from the instructionally mutable: Young children's reasoning about a study of their own design. *Journal of the Learning Sciences, 20*, 50-110.

National Research Council (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice Hall.

Osborne, J., Collins, S., Ratcliffe, M., Millar, R., & Duschl, R. (2003). What "ideas-about-science" should be taught in school science? A Delphi study of the expert community. *Journal of Research in Science Teaching, 40*(7), 692-720.

Penner, D. E., Giles, N. D., Lehrer, R., & Schauble, L. (1997). Building functional models: Designing and elbow. *Journal of Research in Science Teaching, 34*(2), 125-143.

Perry, W. G., Jr. (1970). *Forms of intellectual and ethical development in the college years*. New York: Holt, Rinehart & Winston.

Pintrich, P. R., Marx, R. W., & Boyle, R. A. (1993). Beyond cold conceptual change: the role of motivational beliefs and classroom contextual factors in the process of conceptual change. *Review of Educational Research, 63*(2), 167-199.

Pluta, W. J., Chinn, C. A., & Duncan, R. G. (2011). Learners' epistemic criteria for good scientific models. *Journal of Research in Science Teaching, 48*(5), 486-511.

Roth, W.-M., & Bowen, G. M. (1993). An investigation of problem framing and solving in a grade 8 open-inquiry science program. *Journal of the Learning Sciences, 3*(2), 165-204.

Roth, W.-M., & Desautels, J. (Eds.). (2002). *Science education as/for sociopolitical action*. New York: Peter Lang.

Rudolph, J. L. (2005). Epistemology for the masses: The origins of the "scientific method" in American schools. *History of Education Quarterly, 45*(2), 341-376.

Samarapungavan, A. (1992). Children's judgments in theory choice tasks: Scientific rationality in childhood. *Cognition, 45*, 1-32.

Sandoval, W. A. (2003). Conceptual and epistemic aspects of students' scientific explanations. *Journal of the Learning Sciences, 12*(1), 5-51.

Sandoval, W. A. (2005). Understanding students' practical epistemologies and their influence on learning through inquiry. *Science Education, 89*, 634-656.

Sandoval, W. A., & Reiser, B. J. (2004). Explanation-driven inquiry: Integrating conceptual and epistemic scaffolds for scientific inquiry. *Science Education, 88*, 345-372.

Schwarz, C. V., & White, B. Y. (2005). Metamodeling knowledge: Developing students' understanding of scientific modeling. *Cognition and Instruction, 23*(2), 165-205.

Simon, H. A. (2001). "Seek and ye shall find": How curiosity engenders discovery. In K. Crowley, C. Schunn & T. Okada (Eds.), *Designing for science: Implications from everyday, classroom, and professional settings* (pp. 5-20). Mahwah, NJ: Erlbaum.

Smith, C., Maclin, D., Houghton, C., & Henessey, M. G. (2000). Sixth-grade students' epistemologies of science: The impact of school science experiences on epistemelogical development. *Cognition and Instruction, 18*(3), 349-422.

Solomon, J., Duveen, J., Scot, L., & McCarthy, S. (1992). Teaching about the nature of science through history: Action research in the classroom. *Journal of Research in Science Teaching, 29*, 409-421.

Sternberg, R. J. (2003). *The psychologist's companion: A guide to scientific writing for students and researchers* (4th ed.). Cambridge: Cambridge University Press.

Stokes, D. E. (1997). *Pasteur's quadrant: Basic science and technological innovation*. Washington D.C.: Brookings Institution Press.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

White, B., Collins, A., & Frederiksen, J. (2011). The nature of scientific meta-knowledge. In M. S. Khine & I. M. Saleh (Eds.), *Models and modeling in science education* (Vol. 6, pp. 41-76).

White, B. Y., & Frederiksen, J. R. (1998). Inquiry, modeling, and metacognition: Making science accessible to all students. *Cognition and Instruction, 16*(1), 3-118.

White, B., & Frederiksen, J. (2005). A theoretical framework and approach for fostering metacognitive development. *Educational Psychologist, 40*(4), 211-223.

White, B., Frederiksen, J., & Collins, A. (2009). The interplay of scientific inquiry and metacognition: More than a marriage of convenience. In D. Hacker, J. Dunlosky & A. Graessner (Eds.), *Handbook of Metacognition in Education*: Routledge.

Zion, M., Slezak, M., Shapira, D., Link, E., Bashan, N., Brumer, M., et al. (2004). Dynamic, open inquiry in biology learning. *Science Education, 88*, 728-753.

Appendix A: Interview Protocol

There are three tasks outlined in this interview protocol:

I) Funding Allocation Task
II) Ladybug Task
III) Science Fair Task

Funding Allocation Task

INTERVIEWER**:** *Hello. Thank you for participating in this research study. I am really interested in your ideas about science. I've got a video camera here, so I can remember what you say. Remember that you are a volunteer. That means that you don't need to do this unless you want to. Would you like to answer some questions with me or would your rather not?*

[Wait for response]

[Terminate if child does not actively assent to do the interview].

[If child says he/she would like to proceed:]

INTERVIEWER**:** *Great. I'm going to ask you some questions and there is no right or wrong answer. I just want you to share YOUR ideas. Any time you'd like to stop just tell me and I'll take you back to your classroom.*

Scientists ask questions and they do investigations to try to answer their questions. The investigations may include doing experiments, collecting data or studying something really closely.

To investigate their questions, scientists need to spend money on equipment, materials and to pay their assistants to help the scientist with his/her research. To get funding for their research questions, scientists apply for funding that is given out to scientists that are asking ***worthwhile*** *questions.*

The problem is that there are a lot of scientists and they have a lot of questions that they want to investigate, but there is not enough money to fund all of their research questions.

I want you to pretend that you are the one who decides which scientists' research questions to fund (give money to). Your job is to listen carefully to each of the research questions that the scientists want to investigate, and decide which ***questions do you think would be really worthwhile for SCIENTISTS to research?***

Here is $1,000 in ten $100 bills (fake money).

[Distribute bills to subject.]

INTERVIEWER**:** *There is a set of research questions that scientists want to investigate. For every question that you decide to give money to, you need to give at least $100. You can decide to give more than $100 to a single question. You can distribute the money however you want but you need to distribute ALL of the money.*

I am going to show you each research question one at a time and I will read the question out loud.

I want you to move the question here (point to sign that reads " Probably want to fund") if you think you might want to give money to the scientist so that they can investigate that question. I want you to move the question here (point to sign that reads "Probably don't want to fund") if you think you probably won't give money to that question.

Once I've read all of the questions, and you've moved them into these two piles, I want you to look at the two piles of questions and decide if you want to change your mind about any of them. For example, maybe you put a question in the "probably will fund" but then you change your mind and decide you don't want to give that question any money. Then you can move the question to the "don't fund" pile.

For each question that you decide not to fund, I'm going to ask you to explain why you don't want to fund that question. For each question that you do fund, you are going to place the money next to each question and explain why you gave that much money to that question.

[Interviewer reads one question at a time:

One scientist wants to investigate the question: __________

Subjects move the card with the question to the "probably want to fund" or "probably don't want to fund" piles.]

[Once each question has been read and moved to a pile, Interviewer confirms the current state of the piles.]d

INTERVIEWER: *At the moment, you have [X] number of questions that you said you might want to fund and you have [Y] number of questions that you thought you might not want to fund.*

[Pointing to the "Do not want to fund pile"]

INTERVIEWER: *Before we continue, I want you to look at each of these questions to see if you want to change your mind.* [Read each do-not-fund question out loud]. *Do you want to change your mind about any of these questions? Are there any questions that you* ***do*** *want to fund?*

[Those questions are moved to the "do fund" pile.]

INTERVIEWER: *For each of the remaining questions. Please explain why you decided not to fund this research question.*

[Student explains.]

[Follow up questions to probe student reasoning.] *Why? Why does that mean that you don't want to give money to a scientist to investigate this question?*

[Pointing to the "Do want to fund pile"]

INTERVIEWER: *You have [Y] questions that you want to fund, are any of these questions that you do not want to fund?*

[For each discarded question] INTERVIEWER: *Why did you decide not to fund this question?*

[Follow up questions to probe student reasoning.]

[Pointing to "Fund" pile] INTERVIEWER: *Do you want to fund all of the remaining questions here?*

[Subject is reminded that he/she needs to give each question a minimum of $100 **and that the fewer questions they keep, the more money they can give to each question**.]

[If yes, spread out questions on the table.]

INTERVIEWER: *Now place the amount of money you'd like to give each question right below the question.*

[For each question, ask]: INTERVIEWER: *Why did you give [Z] amount to this question?*

INTERVIEWER: *If you had to give a single question all $1000, which would you fund?*

INTERVIEWER*:* ***In this activity you made decisions about which questions you wanted to fund and which questions you did not want to fund.***
How did you decide?
How did you figure out which questions were worthwhile? What is the difference between the questions that you funded and ones that you didn't fund?

*****

Ladybug Task

INTERVIEWER: *How can plants be protected from pests without using chemicals? What do you think? Do you have any ideas? What ideas do you have?*

INTERVIEWER: *I know a student who is really interested in this problem because he/she knows how bad chemicals are for the people who grow the food, for the people who eat the food and for the environment.*

INTERVIEWER: *He/she wants to do a science investigation to understand what is happening with the pests. He/she thinks that if he/she understands what's happening with the pests, he/she might be able to find a way to protect the plants without using chemicals.*

INTERVIEWER: *When the student went out to the garden to observe, this is what he/she saw.*

[show video of ladybugs eating aphids]

INTERVIEWER: *What do you notice?*

INTERVIEWER: *Those are ladybugs. Ladybugs eat aphids. Aphids are pests that eat broccoli plants.*

INTERVIEWER: *The student also knows that ladybugs live inside flowers.*

[Show photographs of broccoli, ladybugs, aphids and flowers]

Based on what she knows so far, the student has an idea about what MIGHT work to protect the broccoli from aphids.

What do you think his/her idea might be? Explain what his/her idea might be using these pictures if you want.

(If theory not explained clearly): *The student thinks that **maybe** one way of protecting the broccoli plants would be to plant flowers next to the broccoli to attract ladybugs. Then, if the ladybugs come to the flowers, they **might** also eat the aphids for food so that the aphids wouldn't eat all of the broccoli. What do you think?*

INTERVIEWER: *The student wants to do an investigation that will help him/her test his/her idea.*

INTERVIEWER: *Here are a few research questions that the student came up with that he/she thinks might help him/her test her idea that planting flowers near broccoli might attract ladybugs that can eat then eat the aphids (pests). Which question do you think makes the most sense for him/her to research to test his/her ideas?*

Q1 How many aphids can a ladybug eat in one day?
Q2 Does planting flowers next to broccoli plants cause fewer aphids?
Q3 How many spots do ladybugs have on their backs?
Q4 Which flowers do ladybugs like best?

Q5 Which part of broccoli plants do aphids eat first?

[Student explains]

INTERVIEWER: *Are there any other questions that you think might make sense for her to research next?*

Any others?

Are there any questions that you think would NOT make sense to research?

INTERVIEWER: *These are the questions you think make sense to investigate. Assuming the student can only investigate one research question at a time, what order do you think he/she should investigate the questions? Which question should he/she investigate first, second, third etc?*

[Student places the questions in order]

INTERVIEWER: *Please explain why you put the questions in the order that you did.*

*****

Science Fair Task

INTERVIEWER: *There is an* [elementary, middle, high] *school is having a science fair*.

INTERVIEWER*: Do you know what a science fair is?*

[If yes] *What is a science fair?*

[If no] *A science fair is an event where all of the students in grade [X] present the science research projects that they have done.*

INTERVIEWER: *Have you ever been part of a science fair?*

INTERVIEWER: *Months before the science fair, each student in [X] grade has to come up with a question that they want to investigate. Then, each student investigates their question by designing an experiment, collecting data or studying something really closely.*

INTERVIEWER: *At the science fair, there are judges that decide which student has the best science project.*

INTERVIEWER: *In addition to the award for the best project, the judges also give out a special award for the* ***best scientific research question****. The student with the best scientific research question gets the award.*

INTERVIEWER: *The judges are given instructions about how they should choose which scientific question is the best. The instructions are supposed to explain to the judges what makes a scientific research question a really, really good one.*

INTERVIEWER: *If it was up to you to give the judges instructions about how they should decide who gets the award for "best scientific research question", what instructions would you give them? How should the judges decide which science question is the best science question?*

[If student struggles] *For example: The best question award should go to a scientific question that* ________________