GENERATING EPISTEMIC CHANGE THROUGH UTILIZING NATURE OF SCIENCE MATERIAL WITHIN ASTRONOMY

by

Keith Lukas Johnson

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

Doctor of Philosophy

in

Physics

MONTANA STATE UNIVERSITY
Bozeman, Montana

November 2018

ACKNOWLEDGEMENTS

First and foremost I would like to thank my family. Their support and understanding throughout this graduate experience has meant more than they can truly understand. A special thank you to my parents, Jay and Lisa Johnson, for without them I would not be doing what I love with physics today.

I would also like to take time to thank Cooper McCann and Shane Mayer-Gawlik, who would become my closest friends during my time in graduate school. Their kindness, helpfulness, and personalities are something which I will forever cherish.

There may not exist a more fitting advisor for me than Dr. Shannon Willoughby. Her patience and understanding during my development as a physicist and researcher is appreciated beyond what mere words can express. Her thoughtfulness regarding students is something that has made me better as both an individual and an instructor. Special thanks should also go to Dr. Greg Francis, Dr. Mehmet Unver, and Dr. Arthur Bangert, who have always been generous with their time, allowing me to discuss with them a great many things about both research, life, and physics.

I am deeply grateful to all the individuals in the Montana State University Physics department, who have allowed to me opportunities which have had profound impacts on my growth as a physicist. Special thanks here to Margaret Jarrett, Sarah Barutha, Dr. Ives Idzerda, and Dr. John Carlsten.

Finally, I would like to thank Montana Space Grant Consortium for their financial contributions to the research presented within.

TABLE OF CONTENTS

TABLE OF CONTENTS – CONTINUED

TABLE OF CONTENTS – CONTINUED

TABLE OF CONTENTS – CONTINUED

TABLE OF CONTENTS – CONTINUED

LIST OF TABLES

LIST OF TABLES – CONTINUED

Table Page

LIST OF TABLES – CONTINUED

Table Page

LIST OF TABLES – CONTINUED

Table Page

LIST OF TABLES – CONTINUED

Table Page

LIST OF TABLES – CONTINUED

Table Page

LIST OF TABLES – CONTINUED

LIST OF TABLES – CONTINUED

Table Page

LIST OF TABLES – CONTINUED

Table Page

LIST OF TABLES – CONTINUED

LIST OF TABLES – CONTINUED

LIST OF FIGURES

LIST OF FIGURES – CONTINUED

Figure Page

LIST OF FIGURES – CONTINUED

Figure Page

ABSTRACT

At universities across the United States the epistemological beliefs of students have been observed to deteriorate after a semester of instruction in an introductory physics course. What this indicates is that after taking a course in physics most students, for example, may become less capable of delineating between evidence-based reasoning and mere opinion, less likely to believe that they can get better at physics by doing physics, and/or less likely to engage in metacognitive practices. This research proposed that additional course material, based predominantly on the Nature of Science (i.e. the tenets of science) would help prevent this decay of epistemic beliefs. To test this, two years of epistemic data were collected on students in an introductory astronomy class during which time no changes to the course were made. This was subsequently followed by three years of epistemic data collection on students in the same introductory astronomy class in which explicit Nature of Science material had been included. Epistemic data were collected using the Epistemological Belief Assessment for Physical Science (EBAPS).

Results indicated that these course modifications helped to prevent epistemological decay, specifically with respect to student views about science and student beliefs about the role of hard work as compared to natural ability. In order to help identify the impact of the additional material, a complete epistemological framework was identified using student responses to the EBAPS. Functionally, this represents a baseline off which future epistemic work may be conducted within science, as well as outlining a methodology for discovering the complete epistemological framework within students of any study utilizing an epistemic instrument such as the Epistemological Belief Assessment for Physical Science or the Colorado Learning Attitudes about Science Survey. This work also provides further insight into how students are responding to a prominent epistemological instrument within physics, the EBAPS, for which little validation work is present in literature.

INTRODUCTION

The importance of scientific literacy to the general public is perhaps best expressed by Carl Sagan, who once wrote [179]:

> “We live in a society exquisitely dependent on science and technology, in which hardly anyone knows anything about science and technology. This is a clear prescription for disaster.”

This statement by Carl Sagan still holds as much relevance today as it did in 1993. The topic itself has been an area of focus for decades [148]. The National Science Education Standards, which put forth guidelines for the K-12 system, have emphasized the need for students to have a working knowledge of what science is and how science is done [48]. NASA has also expressed their vision on public scientific literacy, stating [6]:

> “Using programmatic tools and resources, the Science Mission Directorate is building strategic educational and public outreach partnerships to enhance the Nations formal education system and contribute to the broad public understanding of science, technology, engineering, and mathematics (STEM).”

When discussing goals for introductory astronomy at the national level, the Education Office of the American Astronomical Society found that the one of the prominent goals of department leaders (e.g. chairs) is to create students who are [159]:

> “[. . .]Interested in and more equipped to follow scientific arguments in the media and more confident of their own critical faculties.”

The American Astronomical Society itself has placed value here, expressing in their mission and vision statement that [2]:

> "The Society promotes broad interest in astronomy, which enhances science literacy and leads many to careers in science and engineering."

This sentiment is mirrored by experts in both astronomy and physics education research [190]:

> "A worthy course goal would be for your [astronomy] students to value the pursuit of science in society as equal to the pursuit of art, ethics, and innovation."

It is with confidence we may state that at both the local and national levels, increasing scientific literacy among the public is a desirable outcome.

The Problem

In 2001, a conservative estimate found that 250,000 undergraduates within the United States and Canada enrolled in an introductory astronomy course and that roughly 10% of college students in all United States colleges are believed to take an introductory astronomy course during their collegiate career [84]. Furthermore, 85% of these individuals taking introductory astronomy are not STEM majors [52]. To summarize, there not only exists a significant portion of the collegiate population which participates in an astronomy class during their undergraduate experience, but also that these same individuals lack notable exposure to science. As many universities do not demand numerous science courses be taken by their non-STEM majors, introductory astronomy represents one of the few opportunities for instructors to instill scientific literacy skills within their students.

In order to improve the scientific literacy of students, these individuals must develop a working knowledge of the Nature of Science itself (NoS). In a broad sense, the nature of science may succinctly be described as the details regarding how science works, which includes all aspects of science: philosophical, sociological, psychological, and historical [142]. Also influential in developing scientific literacy of an individual are their epistemological views. Psychologists tend to define theses epistemologies as pertaining to the nature of knowledge and knowing [104] whereas science educators view epistemologies as associating with the nature of knowledge and knowing within the context of science. Occasionally, these epistemologies within science are treated as being synonymous with NoS [36]. However, within this work, epistemologies will not be treated as being conceptually identical to NoS as defined above. This is largely due to the fact that epistemic instruments within physics go beyond views about knowledge and knowing within science, also including constructs such as beliefs about learning, interest/motivation, and intelligence. Although these additional constructs are also viewed as being context-depend to science they need not directly represent a tenet of science as outlined by [142], but are nevertheless aspects of creating understanding within science. In this respect, we treat scientific literacy as pertaining to views about both the nature of science and epistemological beliefs. This stance appears to also be taken by the National Science Education Standards which define scientific literacy as being [48]:

> “[. . .] greater knowledge and understanding of science subject matter, that is, the knowledge specifically associated with the physical, life, and earth sciences. Scientific literacy also includes understanding the nature of science, the scientific enterprise, and the role of science in society and personal life.”

Overall, complications may occur in attempting to quantify scientific literacy within students simply due to a lack of consensus regarding what NoS is [16].

Measuring an individual's sophistication in views of NoS/epistemology has been largely restricted to forced-choice instruments and, to a lesser extent, interviews. Common instruments claiming to gauge this sophistication include the views of nature of science questionnaire (VNOS, an open-ended survey instrument) as well as the Colorado learning attitudes about science survey (CLASS, a forced-choice likert-style instrument) [10, 128]. As it pertains to this work, we would say that the VNOS questionnaire measures views about NoS as outlined by [142] while the CLASS measures (to varying degrees) views about knowledge, knowing, learning, intelligence, and interest/motivation within physics (any number of which may be influenced by, or closely related to, tenets of NoS).

The importance of scientific literacy and how it may be measured have been discussed in the literature. A review of this literature suggests that an important issue as characterized by Carl Sagan (1993) is that we must generate a more scientifically literate populace so as to keep up with an increasingly more sophisticated technological existence. However, the related problem identified by this research are the efforts to create a more scientifically literate populace at the collegiate level. Specifically, literature has revealed that across the nation student epistemologies in science have been seen to deteriorate after a semester of traditional instruction within a science course [10, 54, 139, 172]. The term "traditional instruction" as discussed in these studies refers to purely lecture-based classrooms with little to no student engagement. Furthermore, these results reported in the literature suggest that from perhaps the only opportunity for non-STEM undergraduates to become more scientifically literate (enrollment in an introductory astronomy class) results in deteriorated beliefs about the nature of science.

Epistemic Beliefs in Introductory Astronomy

The exploration of epistemological beliefs within introductory astronomy remains largely unexplored. One of the few investigations of epistemic beliefs in this domain was conducted by Douglas Duncan of the University of Colorado [60]. The study involved two astronomy-based courses in sequence, the first of which explored fundamental principles (light, gravity, etc.) and applied them to our solar system, and the subsequent course which utilized those same principles outside the solar system. This research effort was motivated by Duncan, who proposed that explicitly taught aspects of NoS (scientific reasoning and judgments, metacognition) may have some positive effect of the epistemic beliefs of students. To this end, the first class in the two-course sequence was left as a control whereas the second class implemented explicit NoS content within the curriculum. In order to measure epistemic differences, Duncan utilized the Epistemological Beliefs Assessment for Physical Science Survey (EBAPS), a 30-item likert-style instrument. The assessment was given to students at the end of the semester for each respective course. It was found that students in the reformed course answered in a significantly different manner than those in the control, and that all the responses were answered more sophisticatedly by the students in the reformed course. Predominantly, these items were associated with beliefs regarding the efficacy of hard work and good study strategies.

Although an intriguing finding, Duncan acknowledged shortcomings within the study design, stating that “We hope that our interesting results will prompt others to try similar experiments in their courses, some of which are more precisely matched” [60]. To elaborate further on the shortcomings, Duncan’s study did not employ true test and control courses, due to the fact that these courses were part of a two-course sequence. The study also utilized two different instructors between the courses. Lastly, students within the second course had one previous semester of astronomy. These shortcomings are important, given the research has found that the epistemic beliefs of the instructor can effect the epistemological state of the students [146, 201]. Previous experience/courses taken by the students have also been influential in epistemologies [9, 183, 184]. Further, the epistemic beliefs of students have been observed to vary in the time between courses, for a two-course sequence [10].

In the fall of 2012 we began what would become a 5-year project designed to further investigate the influence of explicit NoS material incorporated into curricula, and to explore the epistemic beliefs of astronomy students in general.

Research Questions

Question 1

Do basic course modifications in astronomy, structured around explicitly addressing aspects of the nature of science, improve student epistemologies toward science?

This overarching question is at the heart of this epistemological research. From this question, the data gathered would eventually provide a means by which to assess many aspects of the state of student epistemological beliefs within introductory astronomy, beyond just the impact of the NoS implementations. As previously stated, the question itself was motivated by the findings of Douglas Duncan out of the University of Colorado at Boulder [60].

Sub-Question 1a How do students' epistemic beliefs about science change over the course of a semester?

Sub-Question 1b How do students' epistemic beliefs about science change after a semester in which basic course modifications based on the nature of science are implemented?

Question 2

With what epistemological constructs do students rely on within introductory astronomy, with respect to the instrument used to assess these beliefs?

Epistemologies within the realm of physics have been largely treated as being axes of, more-or-less, independent constructs/"beliefs" with each possessing varying degrees of sophistication. Some of the original work regarding this treatment stems from Marlene Schommer in the early 1990s [183–186]. For example, one of the axes Schommer discovered she called "omniscient authority", which represented varying sophistication between two extremes which may be summarized as students being constructors of knowledge through reasoning (most sophisticated) or receivers of knowledge from authority (least sophisticated).

Sub-Question 2a What are the primary epistemological beliefs students rely on within introductory astronomy?

Sub-Question 2b What are the secondary epistemological beliefs students rely on within introductory astronomy?

Research Methodology Synopsis

To assist in answering the research questions above, we collected data on students within an introductory astronomy classroom over the course of five years (2012-2017 academic school years) using the Epistemological Beliefs Assessment for Physical Sciences Survey. In the first 2 years of the study no changes were made to the class, with respect to how it had previously been taught, as its purpose was to function as a baseline. During the final three years specific material focused around the nature of science was added on to class lectures, little to no other content was added or removed. During the entirety of the study the EBAPS was given as a pretest and again as a post-test within the first and last week of class, respectively. Alongside the EBAPS, student written responses to both EBAPS questions and key terminology associated with the EBAPS were acquired. The same instructor was present throughout all the years of the study.

LITERATURE REVIEW

Epistemological Theories

The Nine Stages of Perry's Scheme

Much of the information presented below regarding the work of Perry and his colleagues may be found in text put forth by Perry [165]. In many ways the study of epistemologies began at Harvard with William Perry in the early 1950s. At the time, Perry did not seek to study "epistemologies" as we know them to be today, but instead sought to detail the transformation of college students with respect to their views regarding intellectual development. At the onset of the study, Perry and colleagues expected to be detailing personality traits in the intellectual transformation processes of college students. That is, Perry expected differences in intellectual development to be attributable to differences in personality. This initial focus on personality was due in part to authoritarian personality work present at the time [11]. To study the intellectual differences of college students Perry developed a questionaire known as the "Checklist of Educational Values" (CLEV), which was based on both the aforementioned personality work as well as another instrument called "Stern's Instrument of Beliefs" [192]. The CLEV would first be given as an assessment to a large body of students, followed by more qualitative interview work based on a sample of students from the assessment.

This instrument was administered to 313 freshman during the 1954-1955 academic school year at Harvard. From the 313 freshman, 31 students were invited for open-ended interview work, invitations based predominantly on their answering in extremes of either dualistic or contingent thinking, or at a mean between the two on the CLEV. The interviews themselves typically involved basic questions regarding student development over the academic year, such as "Would you like to say what has stood out for you during the year?" It was through analyzing the interview data that Perry and his colleagues began to see evidence of intellectual differences arising not from personality characteristics, but rather from a developmental process of knowledge. It was from this data that a scheme involving nine stages of intellectual and ethical development was outlined, akin to other stage-based work in psychology being conducted at the time [72,167]. Validation work for the proposed scheme began in 1958. The longitudinal validation study followed 109 freshman from the academic years of 1958-1959 and 1959-1960 throughout their collegiate undergraduate career. In 1970 Perry finalized the scheme which detailed the nine stages through which students were believed to progress in their intellectual development [165]. The progression through these stages is not universal, progression in Perry's scheme can vary based on subject and/or topic. A summary of the stages of this scheme can be found in Table 2.1, along with examples of the nine stages.

Stage 1 Although none of the interviewed students exhibited this trait at any point in their collegiate careers, Perry still proposed this as the fundamental theoretical beginning of development. In this stage all aspects of knowledge are taken as absolute and everything is known (all questions/problems have definitive answers). The student is purely a receiver of knowledge, with any knowledge passed down from authority as being absolute truth which overwrites any pre-held opinions.

Stage 2 Acknowledgement that some authorities may disagree, but there still exists a definitive truth. Reconciliation of competing authorities is often handled utilizing the status of authority figures, such as a teaching assistant as opposed to the instructor. The absence of problems with clear answers within the course may be explained away as the authority figure intentionally creating a vague question and withholding the truth.

Stage 3 Emphasis in this stage is put on how to find the right answers. The only two problems in existance are ones in which the answer is known and ones in which the answer is not yet known. Contradicting information is an indication that there is, currently, no known answer and thus personal opinions are acceptable. Uncertainty is a temporary state of existence and students are focused on deciphering what it is the authority figures want.

Stage 4 There is a separation of sorts within this stage, but both levels of separation acknowledge that uncertainty exists in that there are widespread problems for which authorities hold no absolute truth. On the first level students create a kind of global dualism. Here, there are two kinds of information: Absolute truths from authority figures and subjective knowledge created when figures of authority do not hold the answer. Students on the first level end up on the second level. The second level still dichotomizes the two types of knowledge (authoritative and subjective/opinion-based) but now demands justifications of knowledge from their authoritative sources. In doing so these students also begin to confront their own views with similar scrutiny.

Stage 5 The confrontation of personal opinion within stage four sowed seeds of relativism, where they needed to consider alternative viewpoints and the context in which the knowledge is put forth. In stage five relativistic thinking has become the norm and the realm in which authority figures hold absolute truths has shrunk to the point where it only occurs in special cases. Students are now evaluating most all knowledge and all knowledge must be viewed in context. There may exist multiple answers, but some answers are better than others.

Stage 6 Students in stage six struggle with making commitment to any knowledge and consequently suffer from indecision due to a complete immersion in relativism. As such, they begin to acknowledge the need to commit to a particular solution so as to provide a reference frame in which to establish a knowledge base.

Stages 7-9 Stage seven is established with the individual making a commitment to finding an answer and thus a personal identity, for reasons previously stated. Stage eight has the student begin to consider the implications of their choice in the previous stage and the responsiblity that comes along with it. The final stage involves the individual experiencing the dynamicism that goes along with their commitment. That commitment itself is a continuous, evolving activity and how their stance is influencing personal identity.

Women's Ways of Knowing

One of the most apparent shortcomings of Perry's work was the inherent lack of female perspectives regarding intellectual development. This absence within the educational literature was addressed by Mary Belenky, Blythe Clinchy, Jill Tarule, and Nancy Goldberger culminating in their 1986 text "Women's Ways of Knowing" [20]. It is from this same text that many of the specifics regarding their study are presented below. In order to address the lack of information with respect to the intellectual development of women, Belenky and colleagues decided to conduct their study utilizing only women.

To study the intellectual development of women, a detailed interview was created and administered to 135 women, typically lasting between 2-5 hours. There was a large diversity in the women selected, with a variety of ages (from 16 to 60+), ethnicities, socioeconomic classes, and from numerous types of educational institutions (community college, liberal arts college, etc.). Within the interview

Table 2.1: Perry's Scheme of Intellectual and Ethical Development

Stage	Summary	Example
1	Authority as holder of truth	"The professor is always right"
2	Some authorities are wrong	"The TA is wrong, the professor said [. . .]"
3	Uncertainty exists but is temporary	"My professor doesn't know, but somebody does"
4a	Additional knowledge realm: Opinion	"All TAs have the right to their own opinion"
4b	Justification of knowledge is demanded from Authority	"I asked my professor to show me where they got their information"
5	All knowledge is relativisitic and judged on validity	"I asked both professors to define energy, and both provided unique answers given their fields of study"
6	Complete relativism. A need for orientation is apparant	"I either have to commit to [. . .] or [. . .], otherwise I just don't know where to begin"
7	A commitment is made, a reference point established	"I believe that energy is [. . .]"
8	Implications of commitment explored	"By believing energy to be [. . .] it has led to [. . .]"
9	Commitment seen as an ongoing, evolving activity	"I view energy to be [. . .] and thus my definition of work is a little different. Perhaps I need to reconsider or refine this stance"

were themes surrounding educational experiences, self-concept, moral judgement, relationships, and ways of knowing. The interview questions themselves drew inspiration from other work of the time [89, 122, 165].

An analysis of the interviews led to the realizations that the “ways of knowing” portion of the data had some of the most fundamental perspectives regarding the development of the women. In particular, it became clear to those conducting the study that “women’s epistemological assumptions were central to their perceptions of themselves and their worlds” [20]. From that moment on, the nature of knowing and how knowledge is acquired was at the forefront of the study. Following this philosophy, responses to the interview data were coded with considerations to Perry’s scheme to help assist in their categorization. It was seen that not all responses could be neatly placed into the scheme put forth by Perry and it was decided that a new classification scheme was necessary. This scheme would go on to describe five epistemological perspectives “from which women know and view the world” [20]. These perspectives are presented below and a summary of these perspective, or stages, may be seen in Table 2.2.

Stage 1: Silence Women in this stage believe that they are not capable of understanding or remembering any knowledge being presented to them by an authority. These women also viewed words more as weapons designed to harm and isolate than they do as a means of conveying thoughts or feelings. Effectively, women attempted to stay out of the way as much as possible and have blind obedience toward figures of authority. This is not believed to be a true stage in the progression of epistemological sophistication but rather the result of limited education and poor socioeconomical status. Although not truly akin to the stage of dualism with Perry’s scheme, women in this state do tend to have similar epistemic perspectives.

Stage 2: Received Knowing This stage best fits with Perry's stage of dualism, as women tend to view there as being no ambiguity and believe an absolute truth compliments every question. Furthermore, authorities hold these truths and the student is entirely dependent on these authorities. As such, women struggle to do any original work and feel confused when asked to do so. Words are no longer viewed as weapons but as a way in which to communicate knowledge, although authorities still ultimately are the creators and meaning holders for these words.

Stage 3: Subjectivism No longer do authorities hold absolute truths, in this stage truth is a meaning constructed by the learner. Women go so far as to be suspicious of information provided by authority figures and have a distrust of logic, structured philosophy, and formal investigation. Rather, they look to themselves for knowledge and develop opinions which they acknowledge as being of equal validity as compared to opinions developed by others. Confrontation is avoided so as to preserve relationships with other individuals and although external viewpoints are listened to, they are viewed as not being a reflection of their own personal reality. They have not yet made the realization that "other people are as real as you are" and thus never seek to resolve conflicts of knowledge [22]. Dominating this realm is the value of first-hand experiences, which governs the creation of knowledge from within. In Perry's scheme this perspective, subjectivism, is best associated with multiplicity (stage 3 and 4), where authority figures are beginning to be devalued and opinion holds more influence.

Stage 4: Procedural Knowing This is a stage where knowledge is now viewed as a process, something which needs to be worked at. Not all opinions are viewed as valid anymore and the quality of the knowledge depends on the reliability of the authoritative source. This may best be thought of as a critical thinking perspective, where strategies for evaluating knowledge are developed and relied upon. There are two primary domains with procedural knowing: connected knowing and separate knowing. Connected knowing involves heavy use of adopting the perspective of another individual so as to reconstruct their knowledge through their experiences. This engagement between knowers is motivated through empathy and sustained through intimacy. This is a more submissive domain, in the sense that the individual does not seek to force their knowledge upon the other, but seeks to take the knowledge of the other into consideration by adopting their perspective. Separate knowing is a detachment from all sources of knowledge, including the self. This detachment is also extended toward personal feelings and emotions, relationships do not have a place in the evaluation of knowledge. Consequently, authorities are judged based on the strength of their arguments. Whereas connected knowing may be thought of as an attempt to believe, separate knowing is an attempt to disprove, to doubt. There are no perspectives which are taken and the knowledge seeker challenges the knower. This domain is more agressive and confrontational, but still does not seek to enforce their knowledge upon another.

Stage 5: Constructed Knowing This is essentially the stage of complete understanding of how the process and nature of knowledge is understood. Knowledge is now viewed as something constructed but subject to change due to new experiences, contexts, and perspectives. Knowledge is in a state of construction, destruction, and amending, a fluid state. Constructed knowing aligns strongly with stage 5 of Perry's scheme, where the knower is a crucial part of the knowing process. This also shows strong association with the latter stages of Perry's model, where commitment has been made within relativism and the responsibilities and experiences that accompany that are realized.

Table 2.2: Perspectives Regarding Women's Ways of Knowing

Stage	Summary
Silence	Complete reliance on authority, fly-on-the-wall perspective
Received Knowing	Dualistic philosophy, can reproduce but not create knowledge
Subjectivism	Knowledge comes from personal experience, everybody has right to opinion
Procedural Knowing	Utilize a single approach in determining the validity of knowledge
(a)Connected Knowing	Reliance on adopting alternative perspectives
(b)Separate Knowing	Reliance on objective proof from knower
Constructed Knowing	All knowledge is contextual, utilize subjective and objective approaches to knowing

Epistemological Reflection

Marcia Magolda became interested in long term development of epistemologies during interactions with colleges students and the experiences they had in regards to their changing epistemic views. In particular, how these views changed throughout college and beyond. What would come to be called the "Epistemological Reflection" model, developed by Magolda, focused on these aspects of student development and also gave special attention to the role that gender played during epistemological growth. As per usual, much of the information regarding the methodology and findings associated with the development of the "Epistemological Reflection" model may be found in Magolda's published work on the project [135].

There were two primary phases to the study, both of which utilized interviews as its primary source of information. The initial phase involved following students throughout their collegiate careers with a subsequent phase which continued to monitor epistemic development for nearly a decade after graduation. The study began with 101 participants (51 female, 50 male), all attending a liberal arts college and with a diverse set of majors but with the vast majority being Caucasian. 80 of the original

101 students were able to be followed throughout the entirety of their four (or five, in some instances) years of college. By the end of the study only 39 individuals were left, due to attrition issues involving such things as address changes and or schedule conflicts. All-in-all, the study tracked participants from the age of 18 through to the age of 30, to the best ability of the researcher. Interviews in phase one typically focused on open-ended questions involving the role of their peers or instructors in learning, the nature of knowledge, and their perception of their own work and decision making. During this phase, the Measure of Epistemological Reflections survey was given as well. Phase two (post-college) interviews were typically more conversational in nature, often occurring over a phone and involved the interviewees explaining particularly meaningful learning experiences and the factors surrounding why they were important. All interviews were transcribed and particular themes/patterns were categorized, where handling of interview data essentially followed established protocol [90,193]. Four primary stages of progression were revealed during the analysis of the interview data, with only two of the four occurring predominantly during the college years.

Absolute Knowing The beliefs of absolute knowers can be described with four particular traits. The first is that the role of an instructor was to pass knowledge onto the learner and to make certain the learner received it. Consequently, on the other end, the student (learner) was responsible for acquiring the knowledge from the instructor. As an example, it is the teacher's job to put forth information during class and to check that the students cleanly received the information, whereas the student needed to take responsibility for showing up to class to receive the information. The third trait of an absolute knower is the acknowledgment that peers may enhance your learning experience through accessible explanations and shared knowledge. Lastly, tests of knowledge are a means by which the instructor may gauge student learning. This stage was dominant in the first two years of college, and became negligible after sophomore year.

There are two domains of thought for an absolute knower as well, under the umbrella of traits described previously. The first domain is called "Receiving knowledge" is characterized by a more passive learning style. Individuals within this domain, predominantly females, acquired knowledge mostly through note-taking and listening. They view peers as a means of support in which to confide uncertainties and seek assistance with the material. The second domain, "Mastering knowledge", is majority male who take an active role in their learning, such as asking questions during class or visiting during office hours. The role of their peers was to provide a challenge, relying on things such as interpersonal testing of knowledge (flashcards, quizzing, discussions etc.).

Transitional Knowing This stage was present throughout all of college, prominently so in the junior and senior years. Dichotomy was at the forefront of this stage, where students viewed courses in science and mathematics as having knowledge which was certain and factual, whereas courses in the social sciences and humanities were viewed as having knowledge which needed to be pieced together. Essentially, students simply acquired knowledge in science and mathematics whereas in the social science and humanities students would actively try to understand and apply knowledge.

Much like the previous stage, "Transitional Knowing" also has two sub-categories which are heavily gender-biased. The sub-category "Interpersonal patterns" contained individuals who were more comfortable with expressing opinions, sharing perspectives, and developing personal connections with others. People in this category enjoyed uncertainties in their coursework as the aforementioned traits were best put to use in this context. The majority of individuals belonging to "Interpersonal patterns" were women. At the other end of the spectrum was the category "Impersonal patterns", to which mostly males belonged. Those in this category preferred certainty in the knowledge with which they were working. They isolated themselves in their learning, using their peers as a means by which to reflect upon what they know. These learners relied heavily on logic and typically took a defensive stance when engaging with others to discuss what is known.

Independent Learning At this stage learners have acknowledged that most knowledge is uncertain and that everyone has the right to their own opinions. The role of the instructor or authority figure has become one in which they support the beliefs of the individual while yet trying to develop beneficial views and critical thinking by exploring the context provided. Learners now focus on developing a perspective within the knowledge that they are working, but may still rely on authority to guide questioning and provide challenges. Very few individuals entered this stage during their collegiate career, while over half of those within the study appeared to become independent learners only a single year after graduation. Most often this radical change was a result of the learners being required to become more independent thinking, either due to their job or graduate school.

Separating this stage were behaviors exhibited during interactions with other individuals. Males, generally, were a part of a sub-category called "Individual pattern" learners. In this domain, they struggled to value the views of others while yet placing immense value in views of their own. Females were the majority in the "Interindividual pattern" learners and were more adaptable in their learning because they were willing to consider the views of their peers, yet still valued their own. Overall, both male and females who were seen to belong with the receiving and interpersonal sub-categories previously discussed often showed difficulty in establishing value in their own perspective during this stage.

Contextual Knowing The final stage is one in which effectively no participants in the study were seen to be a part of during their years in college. Individuals in the "Contextual Knowing" stage have now acknowledge that all information is context dependent and to be judged on evidence relevant to the context. These learners are also proficient at identifying the relevant criteria necessary to make informed progression in their knowledge. At this stage many of the gender-biased ways of knowing previously seen are blended together and thus the role gender played all but disappears at this point in life.

The Reflective Judgement Model

The Reflective Judgment (RJ) model put forth by Patricia King and Karen Kitchener is "perhaps the best known and most extensively studied" model of cognitive growth [160]. The model is based on fifteen years of interview data encompassing over seventeen hundred individuals ranging from students in high-school to adults of middle-age. Precise details regarding the extensive development of the reflective judgment model may be found elsewhere in literature [120]. The motivation for the development of the RJ model began with a dissatisfaction of the purely relativistic epistemic description of individuals in the upper levels of Perry's scheme.

King and Kitchener sought to move beyond these relativistic descriptions by thoroughly investigating critical thinking and decision making in the context of ill-structured problems, specifically, to understand the development and deployment of the epistemic constructs which governed reasoning in these situations. In this context, "ill-structured" refers to problems which require considerations of evidence as well as investigation of the source of knowledge, ill-structured problems lack answers of absolute certainty. Examples of "ill-structured" problems can be whether or not violent video games create violent tendencies in those who play them or if there is danger in consuming genetically modified food. Reflective judgments are the result of reflective thinking which is initiated when an individual acknowledges that there exists uncertainty within a problem and that an algorithmic approaches will not yield an absolute truth [53]. The RJ model thus represents primary stages in the development of reflective reasoning leading to reflective judgment. At each stage the epistemic views of individuals are outlined as is their reasoning for making particular judgments, themes between epistemological perspectives and their relationship with these judgments are then connected. King and Kitchener agree that, similar to Perry's scheme, individuals do not necessarily exist strictly in one stage, a philosophy held by others of their time [23, 77]. Rather, students exhibit an approximate range of developmental behavior, as shown in other research involving behavioral development [205].

Pre-Reflective Thinking The first three stages of the RJ model are encompassed by the trait of pre-reflective thinking, students do not yet acknowledge the existence of an ill-structured problem. Stage one is the most primitive stage, where all knowledge is viewed as concrete and any knowledge may be obtained if sought. There is no need for deep justification within this stage because what is observed to be truth is truth and that is all that is required. In stage two knowledge is still absolute but no longer requires only direct observation to be true, knowledge obtained through an authority figure is also seen to be certain. Knowledge may also be temporarily unavailable, but exists with certainty. Beliefs themselves are justified or unjustified through correspondence with authority figures. In the final stage knowledge is again viewed to be absolute but there may now exist temporary uncertainties where it is not known. Beliefs are justified by claiming them as the views of an authority figure or as personal opinion in areas of uncertainty.

Quasi-Reflective Thinking Continuing on through RJ development, stage four is the first in which knowledge is treated as uncertain. The knowledge that an individual carries depends upon a set of variables which create uncertainties and thus contribute to creating a more ambiguous and individualistic, but not necessarily unique, form of knowledge. There is sophistication in judgments as evidence and reasoning are used, however these justifications are idiosyncratic. Knowledge is no longer uncertain in stage five but is instead contextual and subjective. The only thing known are the interpretations made within the context. Justifications are made through the context-dependent evidence and reasoning.

Reflective Thinking Stage six represents a disengagement from personal, subjective interpretation of knowledge and active utilization of other perspectives. As such, knowledge is now constructed from several sources. Beliefs themselves are justified by taking the contextual reasoning and context-dependent evidence from difference perspectives and comparing them against one another. There is also the acknowledgment that other judgments could be equally valid as their own. In the final stage, seven, knowledge is viewed as a continuous process. Knowledge is evaluated and re-evaluated based on reasonability and probability whenever new evidence and/or new perspectives are encountered. Many factors involving reasonability and probability are used when justifying knowledge, claiming they create the most complete understanding possible.

Argumentative Reasoning

Specifics regarding Deanna Kuhn and her study of reasoning may be found in her 1991 text on the matter “The skills of argument”, from which the information below summarizes [125]. Deanna Kuhn initiated a study designed to help elicit an understanding of the informal reasoning people implemented in their daily lives. Much like King and Kitchener, Kuhn wished to understand how people reacted to ill-structured problems they encountered in their everyday lives. In attempting to understand all aspects of argumentative thinking and reasoning, Kuhn came to depend on beliefs individuals held about knowledge and knowing as well.

A unique trait of Kuhn’s research was the wide variety of age and educational groups involved which included 5th-graders, 8th-graders, twelfth graders, undergraduates from a private university, students from a community college, and adults with varying educational and socio-economical backgrounds. In all, forty participants with equal representations of gender and education level were chosen for each of four age groups (teens, 20s, 40s, and 60s). These participants were interviewed twice, with each interview lasting as long as 90 minutes. The interview questions utilized three ill-structured social problems: “What causes prisoners to return to crime after they’re released?”, “What causes children to fail in school?”, and “What causes unemployment?”. Within these questions were continued probing of epistemic perspectives involving multiple viewpoints, certainty of knowledge, and expertise. The interview revealed epistemic beliefs in line with the stage models previously discussed [20, 120, 135, 165].

Kuhn described the epistemological perspectives observed within her data in terms of three categories: Absolutists, Multiplists, and Evaluatist. An absolutist views assertions as being either true or false and knowledge existing in the external world. Critical thinking is then a means by which to connect assertions to the external world. Multiplists have viewpoints akin to subjectivism, opinions are associated with the self and thus knowledge is not solely from the external world. So immersed in opinion are multiplists that they cannot utilize objective means by which to evaluate claims. Critical thinking does not hold much importance in this stage, as all opinions are viewed as equally valid and objectivity is effectively non-existent. Evaluatists now view assertions as judgements which may be subjected to evaluation and scrutiny. Knowledge is still understood as being generated by individuals, but not all assertions are equally valid as evidence and reason are utilized to attempt to determine a more "correct" claim. Essentially this stage has integrated objectivity with subjectivity and critical thinking is now a means by which to make the most accurate conclusions possible.

Epistemological Beliefs

A System of Beliefs

Work by David Schoenfeld in the 1980s provided insight on epistemological beliefs and learning, as did work from Carol Dweck. Dweck and colleagues were the first to establish that children had beliefs about learning and that these beliefs could influence their behavior in various situations, such as what was pursued as a learning goal [63, 64]. It was found that when children are confronted with a difficult task, they are more likely to struggle and become stuck if they believe that learning is a fixed trait. This is contrasted by children who believe that ability to learn is controllable and may change over time, these students were more likely to pursue alternative approaches to the same difficult tasks. Much of Dweck's research in learning would eventually be collected in her 2006 text "Mindset: The New Psychology of Success" [62]. Work by Schoenfeld involved older individuals (high-school) and attempted to understand their beliefs about the nature of mathematics [181, 182]. Schoenfeld found that students who struggled to solve problems often held a kind of dualistic belief that knowledge is passed down from authority. These same students who experienced difficulty with these problems also believed that these problems should take no longer than twelve minutes to solve. Michael Ryan, in 1984, attempted to outline how students with different epistemological beliefs behaved when it came to their comprehension strategies [177]. He found that between particular stages of Perry's scheme there were varying comprehension strategies, opening up the possibility of entire realms of epistemological standards within each. An attempt at collecting all the epistemological aspects of research being done at the time, as well as the difficulty that Ryan had in merging Perry's scheme with metacomprehension, led Marlene Schommer to begin working on the first multidimensional model of epistemologies.

Schommer began her pursuit of a multidimensional model by developing a 63 item questionnaire. The Schommer Epistemological Questionnaire (SEQ) was designed to probe epistemological beliefs through the use of questions such as "Scientists can ultimately get to the truth", and did so utilizing a likert-style response of strongly disagree (1) to strongly agree (5) [183]. Many questions on the instrument took their motivation from previous work, like the CLEV or the work of Dweck and that of Schoenfeld. [64, 165, 180]

The SEQ was administered by Schommer to 266 college students in 1990 (117 junior college and 149 university students) [183]. The junior college students were enrolled in an introductory psychology class while the university students were either enrolled in an introductory educational psychology class or an introductory physics class. Nearly all the students (95%) were freshman or sophomores with an approximately equal amount of females (143) and males (120), where the gender of three students was unknown/unreported. Alongside the SEQ there was administered a survey of student characteristics, so as to simultaneously investigate the relationship of epistemic beliefs and learner characteristics. These characteristics included, but were not limited to: age, gender, year in school, parents' occupation, parents' education level, family composition, ability to make decisions for themselves, and enforcement of rules within family structure. Three educational psychologists reviewed the SEQ results and categorized student responses into 12 subcategories of the original 63 items. Factor analysis was then performed on these 12 subcategory variables, revealing four factors.

Schommer had originally predicted five factors associated with epistemic beliefs: simple knowledge, omniscient authority, certain knowledge, innate ability, and quick learning. Simple knowledge refers to the view of knowledge as a complex network of information as opposed to a collection of facts. Omniscient authority pertains to how students view the acquisition of knowledge as either being handed down by authority or constructed by ones self through reasoning. Innate ability refers back to the work of Dweck and represents how students view the ability to learn as being either innate or dynamic. Certain knowledge explores the tentativeness of knowledge, whether it is set in stone or under constant scrutiny and possible revision. Lastly, quick learning as associated with whether or not students believed learning to occur quickly or not at all. The aforementioned factor analysis in Schommer's 1990 study revealed four roughly independent factors, all of which represented one of the proposed factors with the exception of omniscient authority. Subsequent work was able to replicate these results both with college students and high schoolers [184, 185].

The epistemological sophistication associated with particular factors from Schommer's 1990 study was able to later be linked to particular behavioral traits. Similar to what was seen previously, a sophisticated belief in quick learning (that learning takes time) was found to predict problem solving efficiency in well-structured problems [187]. Sophisticated beliefs in quick learning were also found to correlate with GPA [184, 186]. For ill-structured problems sophisticated beliefs in simple and certain knowledge were seen to relate positively to students' ability to solve these problems [187]. Simple knowledge could also be associated with students' ability to comprehend and study complex text [185].

Epistemological Resources Work on epistemologies thus far has involved either student progression through stages, such as that of Perry, or as the student holding a varying degree of sophistication along partially independent aspects of epistemological beliefs, as detailed by Schommer [165, 183]. David Hammer and Andrew Elby proposed a different perspective for epistemological beliefs which hold a similar theoretical foundation as that of the phenomenological primitive of Andrea diSessa [57]. That is, Hammer and Elby put forth a theoretical account for epistemological views as being the context-dependent activation of more fundamental epistemic resources [95, 96].

Since it's conception in the early 2000s, the epistemological resources theory has been expanded upon but still remains fundamentally unchanged [97, 131, 171]. Examples of epistemic resources include "knowledge as free creation" and "checking". "Knowledge as free creation" is a resource best associated with how individuals come to know something, such as a child whose response to where the idea of for the unique creature they drew came from: "I made it up". "Checking" is a resource that can be involved in answer questions regarding what the individual is doing, this involves justifying knowledge by "checking" with other knowledge acquisition means so as to ensure the correct knowledge is present. These resources are not intended to be concrete (although they could be), but rather have the primary focus of attempting to account for epistemic stances from a more fundamental standpoint. More details and examples regarding these epistemic resources may be found elsewhere in literature

[42, 96]. The idea of epistemological framing has assisted in unifying the purpose of these epistemological resources by providing deeper insight into how they function [97, 171]. Epistemological framing refers to how these resources may turn on or off depending on context within a domain (such as physics) and across domains so as to create a stable epistemological stance [94, 97, 130]. For instance, many novice learners may often view knowledge as handed down from authority and attempt to simply memorize information within a classroom. A student who has had this kind of experience in their high-school chemistry class could come to have the same beliefs in their introductory physics class due to the relatable context (science). In an instance such as this, the theory would claim that these individuals are activating resources such as "knowledge as propagated stuff" (knowledge is passed down from an individual) and "accumulation" (knowledge as a gathered or retrieved collection of information), among possible others. The resources and framework in which this theory is embedded is the subject of ongoing research.

Epistemological Instruments Within Science

Since the work of Schommer in the early 1990's, epistemologies have begun to be more thoroughly investigated within the realm of science. To assist in the study of epistemological views within science, several instruments have been developed which have achieved popularity within the past twenty years. Three instruments in particular stand at the forefront of gauging student epistemologies within science and physics in particular: the Maryland Physics Expectations Survey (MPEX), the Epistemological Beliefs Assessment for Physical Sciences (EBAPS), and the Colorado Learning Attitudes about Science Survey (CLASS) [166]. A fundamental assumption, all of these instruments seek to uncover the internal state of students' beliefs/attitudes by seeing how these individuals respond to a particular statement [83]. As is the case for the instruments listed, this measurement is done using likert-style response options.

MPEX

The work of David Hammer was some of the earliest pioneering work with epistemological beliefs within physics, as his Ph.D. thesis involved a small sample of individuals analyzed in great detail through the use of interviews and careful observation. Not all instructors are capable of studying their students at the extent to which David Hammer had done, the Maryland Physics Expectations Survey was one of the first instruments within physics to attempt to study epistemologies on a large-scale. Specifically, the MPEX was designed to meet the need for assessment of numerous student beliefs within a reasonable (class-length) time period.

Work on the creation of the MPEX began in 1992 at the University of Washington. A test-bed of questions involving the nature of knowledge and knowing within physics was given to an introductory calculus-based physics class. These questions were so chosen as a result of pertinent literature, instrument developer teaching experience, and in discussion with other faculty at Washington. The questions continued to be refined through subsequent implementations of the survey at over 15 other universities. The questions themselves were also validated along the way via discussion with physics education experts, student interview data, and by distributing the survey to other individuals who were proficient in physics. Further details regarding the validation and reliability of the MPEX may be found in a paper by those behind the instrument: Edward Redish, Jeffery M. Saul, and Richard Steinberg [172].

The final MPEX survery utilizes 34 statements asking students to note their feelings regarding each statement by choosing from one of five options on a likert-based scale of options 1-5. For this scale a 1 represents “Strongly Disagree” while a 5 represents “Strongly Agree”. The instrument then quantifies these responses as being either epistemologically “favorable” (4 or 5) or “unfavorable” (1 or 2), with 3 being a neutral option. An example statement within the MPEX is as follows: “The most crucial thing in solving a physics problem is finding the right equation to use”. The assessment, along with more information regarding the MPEX, may be found at its host website [154].

There are six dimensions (mostly orthogonal) which the MPEX assesses, along with an “overall” measure, which is essentially a tally of how students did on all 34 items. These six dimensions are “Independence”, “Coherence”, “Concepts”, “Reality link”, “Math link”, and “Effort”. The former three dimensions (“Independence”, “Coherence”, and “Concepts”) represent more traditionally adhered constructs of epistemology, as pioneered by those previously mentioned. The latter three dimensions (“Reality link”, “Math link”, and “Effort”) were relatively new constructs that the MPEX explores and are not necessarily in-line with what has traditionally been viewed as epistemologies and could perhaps be more appropriately described as “nature of learning” and/or “nature of science” instead. “Independence” assesses the extent to which one constructs their own knowledge as opposed to simply accepting what is presented to them as fact. “Coherence” assesses the extent to which students believe knowledge within physics to be a cohesive framework linking aspects as opposed to just a collection of independent bits of information. “Concepts” is similar to “Coherence” but assesses the extent to which students value the underlying concepts and ideas within physics as opposed to superficial tendencies such as the memorization of how to do a problem and the dependence on formulas. “Reality link” assesses the extent to which student believe that the physics done within the classroom is applicable to their experiences outside of the classroom. “Math link” explicitly asseses the extent to which students believe math to represent physical phenomena. Lastly, “Effort” assesses the extent to which students make use of the information available to them, as opposed to those who do not effectively utilize this information.

EBAPS

The Epistemological Beliefs Assessment for Physical Sciences survey (developed by Andrew Elby, John Frederiksen, Christina Schwarz, and Barbara White) was created to help overcome some issues with other popular epistemic instruments being used at the time, namely, Schommer’s Epistemological Questionnaire (EQ) and the aforementioned MPEX. Schommer’s EQ was operating under the assumption that epistemologies are not necessarily context dependent, that student beliefs do not change with context (such as inter-disciplinary). David Hammer, in work previously outlined, had reason to believe that epistemic stances may change under a variety of different conditions [95, 96]. One such example could be the use of the word “certain”. In psychology a statement about knowledge as “certain” can be viewed quite differently from an identical statement within chemistry [104]. The MPEX was also unsatisfactory in that the instrument held statements which reflected students’ views on how to get good grades, which is not akin to students’ personally held beliefs on knowledge and learning [70]. Hammer’s 1989 work shows several examples for this being a concernable occurrence in response data [93]. Essentially, the authors of the EBAPS wished to create an instrument which, to the best as can be achieved, separated personal beliefs from projected beliefs (for example, one’s own beliefs as opposed to what one believes an instructor might expect) and was fairly robust with respect to mitigating contextual influences.

The EBAPS as it exists today is a 30-item questionnaire which probes various aspects of student beliefs regarding the nature of knowledge, knowing, learning, and intelligence. This survey is similar to the MPEX in that it is a five option response format with responses having varying degrees of epistemological sophistication from “novice” views to “expert” views. There exists three types of items with the EBAPS, categorized into three distinct segments: a likert-style portion, a multiple choice portion, and a debate portion. Items 1-17 on the EBAPS are statements to which the student agrees or disagrees by using a five-option likert-style format ranging from strongly disagree (A) to strongly agree (E) on a discrete A-E scale. Items 18-23 rely on students responding to a question using a set of five possible multiple choice options. These options provide a reasoning alongside an agreement or disagreement with the statement. The final portion of the EBAPS (items 24-30) presents situations to the students in which two individuals are arguing over a particular epistemological stance, generally having one argue a “novice” view while the other argues an “expert” view. Student responses to these debate items are again multiple choice in which the students either agree or disagree to a varying extent with one or both individual/s.

Individuals taking the EBAPS were intended to be upper-level high-school students or lower-level college students, namely, students who are taking introductory physics, chemistry, or physical science courses, it is also claimed to be optimized for algebra-based courses. The assessment itself is recommended to be done via scantron and within 15-22 minutes. Each item on the EBAPS is scored on a scale of 0 to 4, with 4 being the most sophisticated response. The scoring scheme was uniquely developed by the authors in that it is not a linear scale (an “agree” response on one item may not necessarily hold the same value as an “agree” response on another item, similarly responses to the same item may not have equal increments between them). This was done so as to represent the varying levels of sophistication within the responses provided by the students. In essence, items represent a (non-linear) progression from “unsophisticated” to “sophisticated” options. There exists five non-orthogonal dimensions (proposed *a priori* by the authors) to the EBAPS, as well as an overall measure. Scores are obtainable for a given dimension by taking the average the student’s scores along all items pertaining to a dimension. The overall score on the EBAPS is, similarly, the average of all 30 item responses for the student. Items that are left blank by the student are not taken into consideration for these averages.

Authors for the EBAPS claim the instrument to test beliefs along five non-orthogonal dimensions: “Structure of scientific knowledge”, “Nature of knowing and learning”, “Real-life applicability”, “Evolving knowledge”, and “Source of ability to learn”. “Structure of scientific knowledge” quantifies the extent to which students view knowledge within physical science as being either a collection of independent bits (facts and formulas) or as a coherent network of information. “Nature of knowing and learning” quantifies the extent to which students simply take in information without question as opposed to constructing information to create a personalized understanding. The EBAPS examines particular ways in which they might construct their own information such as utilizing metacognition, relating to personal experiences, and consideration of prior knowledge. “Real-life applicability” quantifies the extent to which students believe that activities, concepts, and ideas from within the classroom or laboratory settting apply to the external world. While the EBAPS focuses on trying to tease out personal epistemologies, this dimension takes it a bit further in also attempting to remove any influence regarding whether or not the student desires to apply science to everyday life. “Evolving knowledge” quantifies student views regarding the tentativeness of science. Ideally, a student finds a balance between believing all scientific knowledge to be written in stone and not being able to distinguish opinion from evidence-based argumentation. Lastly, “Source of ability to learn” quantifies the extent to which students believe that learning and doing science is something that individuals are born with as opposed to whether science can be learned through hard work and effort (and good study strategies).

CLASS

The Colorado Learning Attitudes about Science Survey (CLASS) claims to expand upon previous epistemological assessments by probing additional beliefs not typically assessed, such as those put forth within the Schommer Epistemological Questionnaire [183]. The instrument is also heavily contextualized within physics with questions which are open to only a single interpretation, so as to "make the statements as clear and concise as possible and suitable for use in a wide variety of physics courses" [10]. Previous instruments (e.g. MPEX) utilize questions which include two statements that, while answered consistently by experts, are not answered consistently by students.

The CLASS is similar to many other epistemic instruments in that it is a forced-choice, likert-style instrument consisting of axes of beliefs. Specifically, the CLASS uses a five-point likert scale which varies from strongly disagree to strongly agee which, conceputally, varies from least favorable (novice-like) to most favorable (expert-like). The eight axes found on the CLASS are a result of a hybrid approach utilizing both *a priori* expert categories as well as student-based categories defined via factor analytical methods [10]. Students are scored along each of these categories, as well as an overall measrue. Scoring for the class is based around determining a "percent favorable" response rate for each student, i.e. the percentage of responses in which the student has expressed an expert-like view (within a category, and/or overall). The average percent favorable for a category is then determined by taking the average of the individual scores. In this scoring of "percent favorable" responses, the five-point likert scale is collapsed into a three-point likert scale which effectively represents novice-like responses, neutral responses, and expert-like responses. The same approach is taken to determine the percentage of unfavorable responses. Neutral responses are included with neither the favorable or unfavorable percentages and thus represent a unique percentage, that is, the sum of the percentage of favorable, unfavorable, and neutral responses is 100%. A student must have a certain number of responses within a given category to obtain a score for that category and a minimum number of responses on the overall instrument in order to be included in determining an overall score.

The eight scoring categories put forth by the creators of the CLASS are “real world connection”, “personal interest”, “sense making/effort”, “conceptual connections”, “applied conceptual understanding”, “problem solving general”, “problem solving confidence”, and “problem solving sophistication”. The authors themselves do not formally define these axes, stating “one must read the statements to do this.” The interested reader may do so personally, as the items on the CLASS and their categorical associations are readily available in the appendix of the CLASS origin paper [10].

Regardless, a summary of what these categories are believed to explore (as determined by the author of this thesis work) follows. “Real world connection” measures the value of making personal connections of their physics knowledge to the real world (or visa-versa) as well as exploring the applicability of physics to the real world, predominantly the former. “Personal interest” appears to measure a similar construct as “real world connection”, but is better expressed as a measure of the enjoyment a student feels in learning and applying physics. “Sense making” would seem to gauge the extent to which a student desires the construction of a complete set of knowledge. Essentially, sophisticated students in this category will try to connect all the pieces of information available to them into a coherent whole while unsophisticated students will leave gaps in their knowledge. “conceptual connections” represents a category in which students view knowledge in physics as being either isolated or connected, conceptually. Although this category shares much in common with the previous (“sense making”), they are not necessarily the same. For example, students who view physics as being a well-connected body of knowledge need not have the desire to actively seek out these connections. “Applied conceptual understanding” would seem to gauge a student’s ability to apply physics concepts to a variety of scenarios. Unsophisticated responses here likely link to “plug and chug” students who simply look for values and plug them into formulas, often relying on problem mimicry in order to succeed (studying how a specific problem is done, as opposed to studying the concepts behind that problem). “Problem solving confidence” appears to be related to the general belief that there exists an answer to physics problems and that the student can likely find that answer (for the sophisticated students, unsophisticated students would answer in an manner which opposes these views). “Problem solving sophistication” seems heavily intertwined with the ability of a student to assess their own understanding. Unsophisticated students here might believe they understand a physics topic only to find that they do not when tested on their knowledge. It is possible, given the choice of name for the category, that the authors of the CLASS view this category as pertaining to the adaptability of the student (being able to construct knowledge as necessary, based on the scenario). Lastly “problem solving general” combines the latter two categories but also introduces how students treat formulas when learning and doing physics (formulas as tools to be remembered or as manifestations of concept).

Recent factor analysis work indicates that there may only be three primary axes to the CLASS: “personal application and relation to the real world” (a combination of “real world connection” and “personal interest”), “problem solving/learning” ("problem solving sophistication", essentially), and "effort/sense making" (effectively the same as previous discussed) [58].

METHODOLOGY

Introduction

Although the origins of this work are set in exploring the impact that small curricula changes have on student epistemologies within an introductory course in astronomy, it quickly became an exploration of the data collected. That is, the original research questions driving this study have been expanded upon in time. The nature of this culmination of work may now be placed under the umbrella of a single prominent philosophy: What is the current state of student epistemological beliefs within introductory astronomy and how have the changes made to the introductory astronomy course effected these epistemologies?

In attempting to further understand the aforementioned tasks, several distinct areas of focus arose: (1) To determine what the EBAPS indicates about these students' epistemologies and (2) to determine what students' epistemological beliefs are as indicated by student written responses to key aspects of science and specific questions on the EBAPS, respectively. Thus, to help explore student epistemologies within this context a mixed-methods approach was taken in that both qualitative and quantitative data were acquired. The focus of this research also led to us exoploring alternative factor structures within the EBAPS and to, consequently, recommend modifications to items on the EBAPS.

There are three distinct sections within this chapter, designed to specifically address the overarching methodologies of the study. The first will be a discussion of the materials which were added to the course. This section will also present the context for the study by describing the setting, population/s, and instructional style that was present in the classroom throughout. The next section will discuss how the qualitative data (written responses) were analyzed. Finally, the final section will expand upon the mathematics used when quantitative data analyses were conducted.

Astronomy 110

This subsection, titled "Astronomy 110" is an adaptation of previously published material. The following is functionally the same material as that which was originally submitted and accepted to the peer-reviewed *American Journal of Physics* (volume 85, number 6, pages 461-468, co-author/s: Shannon D. Willoughby), with the exception of some basic formatting modifications [212].

Course Details

The Astronomy 110 course is geared toward non-science majors and fulfills a general education science requirement. The demographics of students enrolled in this course are similar to the demographics of the overall student population at the University. At the time of the study, the course had approximately 1000 students enrolled per academic year at a medium-sized land grant institution in the northern United States. Students taking this course were 42.1% male, 39.7% female, and 18.2% did not report their gender. Overall, the average population in the course was 45.5% freshmen, 31.2% sophomores, 12.7% juniors, and 10.6% seniors. The majority of the students are from in-state and, although we did not ask students for their racial identity, the vast majority are Caucasian. (As a whole, the institution consists of 84% self-reported Caucasian students, , 54% are in-state and 54% are male.)

An active learning environment was created during each class meeting, but consistent reinforcement of topics regarding the nature of science were not major parts of the course [85]. The scientific method was introduced during the first week of class, then not directly addressed again. Major topics included naked eye/historical astronomy, cosmology, stellar evolution, and the solar system. In each 200 student section of the course, students formed learning teams each consisting of four students and these learning teams persisted throughout the semester. Students used personal response systems (clickers) each day, lecture tutorials were used during most class periods, and exams consisted of both individual and group portions. Students also completed reflective writing assignments several times during the semester. Outside of class students were asked to read from the text before coming to lecture, but no homework was assigned.

Revised Course

We reveal to you now the results of our baseline data, as revisions to the course were driven by the analysis of this data. After 4 semesters of EBAPS data collection from our unmodified Astronomy 110 course, we found that students were experiencing a decline in epistemic belief structure pertaining to axes two (nature of knowing and learning), four (evolving knowledge), five (source of ability to learn), and the overall measure. The decay in overall student epistemic belief structure is consistent with results across the nation. However, because the EBAPS measures specific areas, or axes, of epistemic beliefs we were able to focus revisions to these areas. Initial changes to the course have been centered on affecting axes two and four, with axis five to be addressed in future work. When the class was revised an active learning environment was still upheld, personal response systems were still used, and learning teams were still employed. The changes implemented involved developing material focused on the nature of science, as well as explicitly relating course material to the nature of science. Any new material was added to the course in lectures that had not previously filled the entire time allotted to each class meeting. Reflective writing and other metacognitive tasks were also included to a larger degree than during the baseline portion of the study, with the goal of positively affecting axis two. We similarly targeted student epistemic beliefs along axis four. Specifically, we addressed issues related to both absolutism and relativism. In practice, these two perils can be thought of as students treating scientific evidence as being set in stone (absolutism) or purely opinion-based (relativism). In order to address issues with absolutism, we discussed a multitude of ways in which understanding in astronomy has changed over the millennia. Classroom discussions combating absolutism include historical changes in solar system models, cosmology, and understanding the scale of the universe. To directly combat relativism, we suggested that often there are clear lines of evidence to suggest why certain things occur. For example, the seasons could be caused by Earths proximity to the Sun, but there is compelling evidence against this. These lines of reasoning are carefully covered in class so students are exposed to the fact that science does indeed change over time (absolutism is not a valid way to consider scientific discoveries), and that sometimes one concept does a superior job of explaining phenomena; not all opinions are equal. In general, we have asked students to search for links between their everyday lives and science, to practice metacognition, to appreciate the tentativeness of science, and to learn in various ways beyond that of simple note taking. Primary course changes are outlined below.

Teaching the Nature of Science Because it affects all aspects of epistemic beliefs (axes two and four included), our primary goal was to make the nature of science an integral part of the course. By regularly discussing how science works and by asking students to see how science is related to their lives, we fundamentally changed our approach to teaching non-science majors. There are several definitions of the nature of science, and we adopt the definition used by Lederman: "[the nature of science] refers to the values and assumptions inherent to the development of scientific knowledge" [127]. Because this topic is fairly broad, we focus specifically on several key aspects: asking students to apply the scientific method, discussing the role of skepticism in scientific discovery, and finding connections to science in their daily lives. Research suggests that the nature of science and how the scientific method is applied needs to be discussed and practiced multiple times in order for students to better understand the concepts and in our revised course, content relating to the nature of science was included at least once each week [15, 127].

We begin on the first day of class by introducing the scientific method in colloquial terms: first, observations are made, then patterns are noticed. An explanation is imagined and from this explanation predictions are made. The loop is then closed as further observations are made to determine whether or not the prediction is consistent with the observation. Eventually, we want to communicate the explanation and predictions to other people so they can also work to determine if the predictions are consistent with their observations. Over the course of a semester the scientific method as defined above is reinforced through a variety of topics, such as moon phases, seasons, and HR diagrams. It is important to focus on each aspect of the scientific method in turn; this is to reduce cognitive overload, or to allow for students to "chunk" the information [145,195]. The goal is to eventually have students understand the scientific method as a coherent whole.

Being skeptical goes hand in hand with applying the scientific method, because once an explanation has been developed for some phenomenon and predictions made, this information can be shared and other people can confirm or deny if the predictions are consistent with their observations. Three to four times during the semester we ask students to complete a worksheet called Science in the Media. Depending on the assignment, students are asked to find an article published by any online source, a major newspaper, or from either a.edu or.gov website. The first Science in the Media assignment is to read and analyze a website claiming the Apollo moon landings were a hoax [46]. The worksheet asks students to consider whether or not personal beliefs are driving the hoax claims, if they think the website is reliable, and whether or not they agree with the conclusions drawn by the author. It also asks students to look for websites that refute the claims, and to determine where the preponderance of evidence points. Worksheet questions were discussed with an information literacy specialist before being implemented, and are included in Appendix E.

All nature of science material in our revised course was aimed to have the greatest effect on both axes two and four, as well as the overall measure of the EBAPS. In this regard, one of our goals in including nature of science material was to share with students the understanding that learning science is a self-constructed process dependent on metacognitive activities. Furthermore, every student has the opportunity to apply science to naked eye astronomy, to practice skepticism, and relate course material to concepts they are most likely already familiar with. Lastly, by acquiring a better understanding of what science is and how it is done we believe our students will find a healthier balance between the perils of absolutism and relativism. (For any specifics or questions regarding course changes and the nature of science, we encourage readers to contact the primary author of this paper.)

Metacognition Of particular interest is the extent to which we can probe students' metacognition, i.e., how they monitor their own thinking and understanding. Metacognition, or "[...] the notion of thinking about one's own thought or thinking process" is a crucial component that enables students to undergo any conceptual change [102, 103, 178, 204]. If a student does not focus on metacognitive activities during the learning process, they are unlikely to overcome any previous misconceptions. Metacognitive actions frequently provide learners with information that influences the selection and use of cognitive strategies [101, 170]. Research suggests that metacognition is a skill that requires practice [141]. To this end we have created several assignments that guide students to consider and analyze their own understanding. These activities include use of exit tickets, exam wrappers, writing their own exam questions, and creating visual representations. A thorough outline of these activities and what they encompass can be found in Appendix E. Assessing metacognition is no trivial task, as it is not directly observable and often involves many types of knowledge. Examples of valid instruments assessing solely metacognition include the Learning and Study Strategies Instrument (LASSI) and the Metacognitive Awareness Inventory. Both of these instruments consist of over items, and only test metacognition [12]. Axis two of the EBAPS allows us to measure, and therefore monitor, student metacognition.

End of previously published material as adapted from [212].

Qualitative Data Analysis

Written Response Data

Written responses to several EBAPS questions and to some key terminology within the EBAPS were acquired throughout the course of the study. The manner in which these responses were collected was through the use of an online homework system. Typically, the questions that were posed to the students were motivated by findings within statistical analyses done on the student EBAPS data around the same time. Hence, not every EBAPS questions was posed to the students and neither were these questions necessarily posed in the same order as is done on the EBAPS. By the end of the study we were able to obtain written student responses to 24 of the 30 items on the EBAPS. Most of the questions were posed to the students in our introductory astronomy class during the final two weeks of a semester.

The first online homework system utilized was Sapling Learning. Student responses to the questions were always volunteer-based and involved the following introductory prompt:

> "Answer the following questions below to the best of your ability. If you would like to know more regarding the nature of these questions then continue reading the section immediately below, otherwise you may begin answering.
>
> In regards to student epistemologies research has shown that non-science majors who take general education science courses have degraded attitudes toward science upon completion of these courses. Epistemological beliefs can be defined as beliefs about what it means to learn and how knowledge is constructed. The following optional response questions are designed

> to help assist us in determining the development and/or persistence of particular beliefs that individuals hold regarding science."

These optional written responses were assigned alongside a mandatory online homework assignment. The written responses would often number between 50-100 student volunteers. This manner of acquiring written responses to the EBAPS questions was done in spring of 2016 (EBAPS questions 5, 9, 16, and 22), spring of 2017 (EBAPS questions 1, 6, 10, 12, 24, 29, 30), and fall of 2017 (EBAPS questions 3, 8, 11, 14, 15, 17, and 28). During the spring of 2018 a different online homework system was used, Pearson Mastering Astronomy. This homework system behaved similarly to Sapling Learning in that EBAPS questions would be put forth alongside a required online homework assignment. The voluntary written responses were again posed within the final two weeks of the semester, however, due to the grading system Pearson utilizes the written responses were given as extra credit (if all responses were completed, the students would have their lowest online homework grade replaced with full credit). Likely due to the promise of extra credit, spring 2018 saw between 100-200 written responses to the posed questions (EBAPS questions 4, 7, 21, 26, and 27). On Pearson, the questions were posed with an introduction as follows:

> "Answer the following five questions below to the best of your ability. If you complete these items, they will replace your lowest homework score with a full-credit score (10/10). If you would like to know more regarding the nature of these questions then continue reading the section immediately below, otherwise you may begin answering.
>
> Additional Info (If Interested): In regards to student epistemologies research has shown that non-science majors who take general education science courses have degraded attitudes toward science upon completion

of these courses. Epistemological beliefs can be defined as beliefs about what it means to learn and how knowledge is constructed. The following response questions are designed to help assist us in determining the development and/or persistence of particular beliefs that individuals hold regarding science and/or life."

The responses were analyzed utilizing predominantly a two-step method, having many similarities with the constant comparison method [90, 197]. The leading philosophy guiding the analysis of the written responses was to let the students dictate the coding. As the EBAPS questions are intrinsically epistemological, the primary focus was on determining student reasoning behind their responses. To this end we hoped to find constructs/beliefs students relied on when responding to contexts designed to engage their personal epistemic beliefs.

The first step of the written response analysis involved what is referred to as "initial coding" or "open" coding [34, 45, 90]. Initial coding involves the partitioning of response data into distinct pieces such as key concepts, ideas, words, etc. The researcher need not necessarily be looking for a particular trait as they begin coding, instead letting the data speak for itself. Often times in this coding process new trends are observed as the data is parsed, hence codes are often created throughout. These codes need not be concrete, as not only can they change as the data is parsed but may also be altered or abandoned in subsequent passes of the data. This type of coding methodology fits well with our philosophical approach, described above, in that it provides a means by which to approach the data with an open-mind and helps avoids pre-conceived epistemic notions. During this step, the codes that were generated focused on what was perceived to be the main argumentative focus within student responses. Formally, during this "initial coding", a methodology known as "topic coding" or "descriptive coding" was most frequently employed [147, 213]. As Tesch best expresses, topic coding utilizes codes which are "identifications of the topic, not abbreviations of the content" [198]. Here, Tesch views the "topic" as the general theme of the statement whereas the "content" refers to the pieces of the statement which generate the final perspective of the individual.

The second step during the analysis of the written response data was the constant comparison and refinement of the codes that existed. Although similar in many regards to the first step, this step was geared less towards developing codes and more towards fitting and grouping the codes that existed to a common theme/category. This often involved looking at particular codes from a broader perspective. Formally, this step is what researchers would claim to be "axial coding" [34, 90]. Conceptually, axial coding is named such because you are taking your codes and attempting to align them along a particular thematic axis/dimension. There is no need for a theoretical/predetermined axis here, as one is simply attempting to identify a common theme. Furthermore, one does not seek to align all codes along a single axis, this framework is allowed to be multidimensional.

Typically the coding process may end with what is called "theoretical coding" [45]. This stage is similar to the previous, yet distinct in that you are taking the axes/categories that existed in the previous step and collapsing them down into a single explanatory dimension. This is not to say that the categories are no longer distinct, but rather that they have now all been connected to a single common theme. This stage was not an explicit part of our coding process for two primary reasons. The first being that the framework in which these questions are embedded are already going to resemble epistemic beliefs, that the backbone connecting all these categories is the epistemic belief structure of the students. Epistemic belief structures themselves within the context of the study and throughout physics education research are viewed largely to be made of relatively distinct dimensions. As such, our philosophical approach would resemble such. The second reason for the exclusion of this coding stage is that as we discuss the prominent response axes found within these analyses, we often elaborate on their inter-connectivity and the role they play in influencing the student responses and are hence inevitably touching on the "theoretical coding" stage via a written elaboration as opposed to an explicit coding.

The coding steps outlined above are true for all written response data outside of the data collected from spring 2018. Responses from spring 2018 did not include step two, in which all codes were assigned to a generic category. Instead, spring 2018 written responses utilized a more thorough account of "initial coding" from step one. As opposed to the "descriptive coding" previously used, more emphasis was now placed on the contents of the response. Multiple codes would now be assigned to portions of a student's response. This was found to be more enlightening as often a student would have multiple reasons for responding in the manner they did (largely due to how the EBAPS questions are constructed). Of course, the general trend of "initial coding" was still followed, that being the tentativeness and recursive refinement of the codes. As pieces of student responses were now being emphasized, "values coding" was more frequently at use [87, 126]. In values coding, more focus is put on how an individual thinks and feels about a topic. In general, the codes in values coding align with student attitudes and beliefs with respect to the question being addressed.

The previous two-step approach was abandoned for the spring 2018 response data due to time constraints. However, generating multiple codes within a single student response was valuable from a theoretical perspective, as one might often expect that multiple epistemic beliefs to be functioning at once when responding to an EBAPS question [69]. This change in initial coding allowed more insight into student reasoning and as such all other response data was coded again (previous coding results were still retained and are discussed) utilizing the same approach as was done with the spring 2018 data.

As the results of the written response analyses are discussed below, they will be presented in a particular order so as to retain clarity due to the use of multiple coding schemes. First, the EBAPS question as posed to the student will be presented, followed by any concerns regarding the manner in which students responded to the question. Note that some questions were altered slightly from their EBAPS counterpart, typically done to better contextualize the scenario (name changes, mentions of astronomy, etc.). In later years, modification of any EBAPS questions was dropped, as it was believed that exact preservation of the question was important. Second, the newest coding scheme (the spring 2018 scheme) will be discussed as it provides more intricate details regarding the nature of the student responses. Then, at last, the previously outlined two-step scheme will be presented as a means by which to summarize the bulk of the findings regarding a given question. For additional clarity, a conscious effort will be made to discuss the revised methodology as using discrete "codes" whereas results from the two-step methodology (where applicable) will be referred to as the more general "themes".

Quantitative Data

This section functions as an introduction to the variety of statistical tests utilized throughout this work. Common quantifiable measures of data behavior typically involve determining whether or not some assumed model provides a good fit for the sample data in question or if two parameters which define a distribution are the same value. The subsections below begin with a general discussion of hypothesis testing, which relies on the assumption that the data in a sample are drawn from a well-defined distribution. In the context of this study, hypothesis testing directly pertains to several test types: Bartlett's test of sphericity, independent t-tests, z-tests of proportions, and F-tests. All the aforementioned tests fit assumptions about the distributions from which they are believed to have been drawn. For data which do not fit all (or most) of the distributional assumptions, non-parametric methods are utilized: the chi-squared goodness-of-fit test, chi-square test of homogeneity, Wilcoxin signed-rank test, and the Mann-Whitney U test are those utilized in this study. Although non-parametric tests are designed for data which are not ideal (violate distributional assumptions), they often simply transform the data into a statistic which has a well-defined distribution and hence can be subjected to hypothesis testing. Other than significance testing, the most prominent work with processing data involves exploratory factor analysis, which provides a means by which to analyze student epistemological behavior based on their responses to the EBAPS.

Hypothesis Testing

Introduction Many of the statistical tests performed within this work depend on observing a specific value from within a distribution and determining the probability of having observed such a value. This section on hypothesis testing presents the basic formulation of how these tests are performed.

Mathematical Foundation We begin by considering that for any sampling of data (represented by a vector of values, $\vec{x}$) a scalar test statistic called $\gamma(\vec{x})$ may be constructed. Depending on criteria set by the researcher, this test statistic may turn in to a "critical value" for use in statistical tests of significance. For example, should a distribution behave in such a manner that all values x, of $\vec{x}$, obey $\infty < x < \infty$ then we may determine that the probability of observing an x value greater than a critical value γ_{crit} given some null hypothesis H_o is:

$$P(x > \gamma_{crit}|H_o) = \int_{\gamma_{crit}}^{\infty} P(x|H_o)dx = \alpha \tag{3.1}$$

Here, α represents the statistical certainty that we desire, i.e. how certain do we wish to be that we observe x given H_o.

Before discussing how these test statistics may be acquired, we must first discuss more fundamental constructs. Consider n random data samples, symbolized by $\vec{x} = \{x_1, x_2, \ldots, x_n\}$, drawn from a distribution with probability density function $f(x|\theta)$ where $\vec{\theta}$ represents some parameter/s which specify the distribution. It follows that a likelihood function may be defined as:

$$L(\vec{\theta}|\vec{x}) = f(x_1|\vec{\theta})f(x_2|\vec{\theta})\ldots f(x_n|\vec{\theta}) \tag{3.2}$$

The likelihood function, then, can be used in determining the most likely value/s of $\vec{\theta}$ given $\vec{x}$ (typically through maximum likelihood methods). A generalized likelihood ratio [174] gives way to the test statistic via:

$$\gamma(\vec{x}) = \frac{L(\hat{\theta}_S|\vec{x})}{L(\hat{\theta}|\vec{x})} \tag{3.3}$$

Where $\hat{\theta}$ is the maximum likelihood estimate/s for the population parameter value/s and $\hat{\theta}_S$ is the maximum likelihood estimate/s for the sample subspace, dictated by the null hypothesis. Uncovering the critical statistic is thus unique to the sampling distribution and researcher input. This begins by the generalized likelihood ratio test:

$$\gamma(\vec{x}) \leq k \tag{3.4}$$

Here, k is some arbitrary constant. Should γ behave as outlined above, then the null hypothesis is to be rejected. This holds true for any function $f(\gamma)$ as well, so long as it is a monotonically increasing function of γ [174].

As an example, consider taking a set of n independent samples from a Gaussian distribution with a known standard deviation σ_o but an unknown mean μ. We wish to test the null hypothesis that $\mu = 0$ at the 95% certainty level. The probability density function for each sample x_i would be:

$$f(x_i) = \frac{1}{\sqrt{2\pi\sigma_o^2}} \exp\left(-\frac{(x_i-\mu)^2}{2\sigma_o^2}\right) \tag{3.5}$$

Constructing the likelihood function would have:

$$L(\sigma_o, \mu|\vec{x}) = f(x_1|\sigma_o, \mu) f(x_2|\sigma_o, \mu) \dots f(x_n|\sigma_o, \mu) \tag{3.6}$$

Which yields:

$$L(\sigma_o, \mu|\vec{x}) = \frac{1}{(2\pi\sigma_o^2)^{\frac{n}{2}}} \exp\left(-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma_o^2}\right) \tag{3.7}$$

It follows that the test statistic be:

$$\gamma(\vec{x}) = \frac{L(\sigma_o, \mu=0|\vec{x})}{L(\sigma_o, \bar{x}|\vec{x})} = \frac{\exp\left(-\frac{1}{2}\sum_{i=1}^{n} x_i^2\right)}{\exp\left(-\frac{1}{2}\sum_{i=1}^{n}(x_i-\bar{x})^2\right)} = \exp\left(-\frac{1}{2}n\bar{x}^2\right) \tag{3.8}$$

In the likelihood ratio used above, the values of σ^2 and μ which maximize $L(\sigma, \mu|\vec{x})$ in a population are s^2 (σ_o^2, in this example) and $\bar{x}$, respectively (s^2 is the sample standard deviation, $\bar{x}$ is the sample mean) [174]. To then find the critical value γ_{crit}:

$$\int_{-\infty}^{\gamma_{crit}} P(x|H_o)dx = 0.05 \tag{3.9}$$

In the integral it is recommended to let $\nu = -2\ln(\gamma) = n\bar{x}^2$, and proceed from here [174]. As ν is a monotonically decreasing function with respect to γ, the new integral form is:

$$\int_{\nu_{crit}}^{\infty} P(\nu|H_o)d\nu = 0.05 \tag{3.10}$$

Upon determining the value of ν_{crit}, we relate back to the sampling parameter $n\bar{x}^2$ and find that the critical values for $\bar{x}$ be of the form $\bar{x}_{crit} > |\frac{\nu_{crit}}{n}|$. Thus we reject the null hypothesis if the mean of our n sampled values is such that $\bar{x} > \frac{\nu_{crit}}{n}$ or $\bar{x} < -\frac{\nu_{crit}}{n}$.

Exploratory Factor Analysis

Introduction Exploratory factor analysis (EFA) is the utilization of a particular pattern finding technique to discover underlying constructs which may be responsible for influencing (or linking) data [51]. These mathematical techniques find their use within a variety of fields such as the social sciences, geology, and economics.

For a theoretical example of how EFA could be employed, consider the measurement of average yearly surface temperatures at several locations (15 locations, say) across the globe. Furthermore, these measurements will be said to have been taken each year over the course of 150 years. In this example, the 15 locations will function as the columns within a data matrix and the average yearly surface temperature measurements (150 per city) function as the rows of the same data matrix. An exploratory factor analysis on this data should result in the grouping of one or more sets of locations. It would then be up to the researcher to provide plausible reasoning for the resulting group/s. One might very well find that these cities appear to be grouped by biome (tundra, temperate, desert, tropical, etc.). That being the case, EFA in this example could have identified longitudinal location on earth as being the underlying construct responsible for the differences seen in average yearly temperature. Herein lies one of the shortcomings of EFA, that the data is complex and depends on far more than just what is observed [32, 202]. The major factors discovered within an analysis are not necessarily the only factors, but may just be the most prominent factors. In the example above, longitudinal location may have been picked up within EFA, but not factors such as altitude or ocean proximity, which should also play a significant role in varying surface temperature measurements.

The theory and mathematical origins of EFA was introduced in 1904 by Spearman for the identification of a single factor, followed by a generalized (multi-factor) approach put forth by Thurstone in 1947 [191, 199]. Mathematically, EFA utilizes correlations (or covariances) among items to identify a factor or set of factors whose items share a common variance (i.e. factors) [112]. Conceptually, the intent of EFA is to discover these underlying factors utilizing parsimony [98]. Within EFA, parsimony represents the ability to identify relationships between as many observed variables as possible (all, ideally) in the simplest manner possible (i.e. with the fewest factors) [158].

Mathematical Foundation Specific mathematical details of both the original one-factor model and the extended multi-factor model can be found in their origin papers by Spearman and by Thurstone, respectively [191, 199]. To understand exploratory factor analysis (EFA), it may be best to understand what principal components analysis (PCA) is first [107, 108, 163].

As it would relate to the calculations done within this study, the general philosophy behind PCA is to apply eigenvalue decomposition to a $N \times M$ correlation matrix $\widetilde{\mathbf{R}}$. The principal components of $\widetilde{\mathbf{R}}$ are then simply M orthogonal axes which attempt to explain as much variation within the data as possible. These axes would correspond to the axes of an ellipsoid in M dimensional space.

Consider a correlation matrix, $\widetilde{\mathbf{R}}$, such that there exist principal components, $\widetilde{\mathbf{C}}$, in which

$$\widetilde{\mathbf{R}} = \widetilde{\mathbf{C}}\widetilde{\mathbf{C}}^T \tag{3.11}$$

If $\widetilde{\mathbf{X}}$ is a matrix of orthogonal eigenvectors of $\widetilde{\mathbf{R}}$, then

$$\widetilde{\mathbf{R}}\widetilde{\mathbf{X}} = \widetilde{\mathbf{X}}\lambda \tag{3.12}$$

where λ is a diagonal matrix of eigenvalues of $\widetilde{\mathbf{X}}$. It follows that if

$$\widetilde{\mathbf{R}} = \widetilde{\mathbf{X}}\lambda\widetilde{\mathbf{X}}^T \tag{3.13}$$

then $\widetilde{\mathbf{C}}$ is

$$\widetilde{\mathbf{C}} = \widetilde{\mathbf{X}}\sqrt{\lambda} \tag{3.14}$$

$\widetilde{\mathbf{C}}$, then, is just a matrix whose columns are eigenvectors scaled by the square root of the eigenvalues. If $\widetilde{\mathbf{C}}$ were to contain all the principal components of $\widetilde{\mathbf{R}}$ then $\widetilde{\mathbf{C}}$ would be perfectly replicated by $\widetilde{\mathbf{C}}\widetilde{\mathbf{C}}^T$.

The goal of PCA is not perfect replication of $\widetilde{\mathbf{R}}$ but rather to achieve a parsimonious representation of $\widetilde{\mathbf{R}}$, i.e. to represent $\widetilde{\mathbf{R}}$ accurately with as few principal components as possible. As such, only a predetermined number of components will be retained, with a variety of options available for doing so [49,153]. Nevertheless, when fewer than M principal components are retained then $\widetilde{\mathbf{R}} \neq \widetilde{\mathbf{C}}\widetilde{\mathbf{C}}^T$ and the residue matrix, $\widetilde{\mathbf{R}}^*$, will be non-zero

$$\widetilde{\mathbf{R}}^* = \widetilde{\mathbf{R}} - \widetilde{\mathbf{C}}\widetilde{\mathbf{C}}^T \tag{3.15}$$

As mentioned previously, PCA is an attempt at reducing a set a variables into a smaller set of descriptive components. In general, PCA is constructed around using observed variables, x_j, and weighting coefficients, w_{ij}, to generate a score which pertains to a particular principal component

$$Y_i = \sum_{j=1}^{M} w_{ij}x_j \tag{3.16}$$

Within PCA, the first principal component attempts to explain as much of the variation in the data as possible. Subsequent components attempt to account for the maximal amount of variance that is left over after the previous principal component has been accounted for. In geometric terms, orientation of the axes of the aforementioned ellipsoid are related to the eigenvectors and hence the linear combination of weights. The corresponding eigenvalues represent the length of each ellipsoidal axis and are related to the amount of variance that is accounted for. This variance of the items within $\widetilde{\mathbf{R}}$ may be accounted for as follows

$$\sigma^2 = \sum_{j=1}^{M} \lambda_j \tag{3.17}$$

As this presentation of PCA was done with a correlation matrix, $\widetilde{\mathbf{R}}$, the total variance should be equivalent to the total number of principal components, M.

Consider a measurement matrix $\widetilde{\mathbf{V}}_{ij}$ whose entries represent the ith student score on the jth item. In measurement theory it would follow that

$$\widetilde{\mathbf{V}}_{ij} = \widetilde{\mathbf{T}}_{ij} + \widetilde{\mathbf{E}}_{ij} \tag{3.18}$$

where $\widetilde{\mathbf{T}}_{ij}$ are the true scores a student achieved and $\widetilde{\mathbf{E}}_{ij}$ is the measurement error associated with their actual score. Subsequently, the variance of the observed scores would be

$$\sigma_V^2 = \sigma_T^2 + \sigma_E^2 \tag{3.19}$$

Within PCA, the entirety of the measurement variance, σ_V^2, is taken into account (technically, PCA makes the assumption that $\sigma_E^2 = 0$ and would thus be said to theoretically only consider σ_T^2). Exploratory factor analysis attempts to expand the true variance (σ_T^2) even further into a "common" (σ_C^2) and "specific" (σ_S^2) variance, that is

$$\sigma_T^2 = \sigma_C^2 + \sigma_S^2 \tag{3.20}$$

Where σ_C^2 represents variance that is unique to a set of factors (not principal components, but similar) and σ_S^2 represents a variance that is not explained by either the factors (potentially due to the content of a given item) or by measurement error. σ_S^2 and σ_E^2 are grouped together as a single variance called the "unique" variance ($\sigma_U^2 = \sigma_S^2 + \sigma_E^2$), thus giving factor analysis the overall model variance representation of

$$\sigma_V^2 = \sigma_C^2 + \sigma_U^2 \tag{3.21}$$

Within EFA, σ_C^2 are referred to as "communalities" (often shown as h^2) and σ_U^2 are referred to as "uniquenesses" (often shown as u^2).

While PCA may be viewed as a type of dimension-reduction for which the scores on components may be calculated exactly, scores on factors within EFA are estimates. This difference, of course, has its basis in the methodology of EFA and the emphasis put on accounting for common variance. The approach of EFA is not all that different than PCA, in that $\widetilde{\mathbf{R}}$ is still the matrix attempting to be modeled. Unlike PCA, however, $\widetilde{\mathbf{R}}$ is now modeled as

$$\widetilde{\mathbf{R}} \approx \widetilde{\mathbf{F}}\widetilde{\mathbf{F}}^T + \widetilde{\mathbf{U}}^2 \tag{3.22}$$

where $\widetilde{\mathbf{U}}^2$ is a diagonal matrix of uniquenesses and $\widetilde{\mathbf{F}}$ is a matrix whose column vectors are the common factors. The philosophy behind EFA is to construct factors which are only responsible for having created the common error within the data.

The general process behind EFA as performed in this study may be outlined as follows:

1. Obtain an initial estimate of $\widetilde{\mathbf{U}}^2$ (This is optional, matrix of zeroes otherwise)

2. Obtain a set predetermined number of principal components (eigenvalues and

eigenvectors) for $\widetilde{\mathbf{R}} - \widetilde{\mathbf{U}}^2$

3. Acquire the residual matrix $\widetilde{\mathbf{R}}^* = \widetilde{\mathbf{R}} - \widetilde{\mathbf{F}}\widetilde{\mathbf{F}}^T$, where $\widetilde{\mathbf{F}}$ is the product of your eigenvectors and eigenvalues (principal components)

4. Find the new matrix of uniquenesses $\widetilde{\mathbf{U}}^2 = diag(\widetilde{\mathbf{R}}^*)$

5. Repeat steps 2-4 until the difference in the value of $\mathbf{I} - \widetilde{\mathbf{U}}^2$ from one iteration to the next is arbitrarily small

The intial estimates for $\widetilde{\mathbf{U}}^2$ mentioned in step 1 are done to make the problem less computationally expensive and may be obtained via

$$\mathbf{U}^2 = \mathbf{I} - [diag(\widetilde{\mathbf{R}}^{-1})]^{-1} \tag{3.23}$$

The estimate above (provided by Harris in 1978) will put a lower bound on the communalities, other techniques do exist for this initial estimates for the communalities, however [99,214]. Thus the eigenvectors in EFA are constructed based on items whose variance is purely that which is due to "common" factors [175]. The factors in the equation $\widetilde{\mathbf{R}} \approx \widetilde{\mathbf{F}}\widetilde{\mathbf{F}}^T + \widetilde{\mathbf{U}}^2$ may be estimated through means other than just the iterative method outlined above (most commonly known as principal axis factoring). Other popular methods include ordinary least squares, generalized least squares, and maximum likelihood [13, 123, 211].

Following factor extraction in EFA (and PCA) comes the choice of whether or not to rotate these factor solutions. Before rotation, the factors are all orthogonal (as they are eigenvectors). As such, these factors describe the variance within the data in the most efficient manner possible but in a manner which is not always clearly interpretative. Rotations themselves will occur in "loading" space, where a "loading" refers to the correlation between an item and a factor. If one plots the loading values of all your items from one factor versus the loading values of all your items from a separate factor, the resulting plot would represent a "loading" space. The axes in this space represent the corresponding eigenvectors of your chosen factors. These axes may then be rotated in "loading" space to better align with the distribution of items. In PCA the unrotated item loadings are the elements of the principal components themselves $\widetilde{\mathbf{X}}\sqrt{\lambda}$. Similarly for EFA, except that the eigenvector/eigenvalues now pertain to those acquired from $\widetilde{\mathbf{R}} - \widetilde{\mathbf{U}}^2$.

Orthogonal rotation of the axes (eigenvectors) are typically done when there is no reason to believe that the axes represent constructs which are correlated, these types of rotations are utilized within PCA but may be employed within EFA as well [51]. Two common types of orthogonal rotations are varimax and quartimax. The varimax algorithm is based on obtaining a factor/component structure which has a single item being strongly associated with the given factor/component, i.e. it attempts to minimize the number of items with strong correlations (loadings) on factors/components [73]. Quartimax attempts to maximize the loadings of all items onto a single component/factor. It also strives to have these same items load high onto a second factor, while keeping loadings onto any other factors at a minimum [151]. In this regard, the philosophy behind quartimax is that all items are explained well by a single prominent factor and also by secondary factors which explain a common trait attributed to a subset of the items. Upon an orthogonal rotation, the original components/factors simply change as $\widetilde{\mathbf{C}} \rightarrow \widetilde{\mathbf{C}}\widetilde{\mathbf{T}}$, where $\widetilde{\mathbf{T}}$ is the transformation matrix used in rotating the original factors.

Oblique rotations remove the restriction that the axes must be orthogonal to each other and are unique to EFA. This type of rotation is employed when one expects that there be a relationship between the factors, as is usually the case in practice [75, 176]. The direct oblimin and promax methods of oblique rotations are two which are often put into practice. Oblimin techniques start with an unrotated solution and then attempt to achieve simple structure while utilizing a parameter which allows for a user-specified amount of allowable correlation between factors [114, 199]. These parameters typically range from -4 (representing an orthogonal rotation) to 1 (representing the most highly correlated factors possible). When employing oblimin techniques within data, the default parameter for both R and SPSS computational software is 0 (mild correlations allowable). Promax rotations will start with an orthogonal varimax rotation, then attempt deviations from there. This is typically done by manipulation of the loadings to higher powers until simple structure is achieved [65].

Upon any rotation, the final approximation to the original $\widetilde{\mathbf{R}}$ matrix is now

$$\widetilde{\mathbf{R}} \approx \widetilde{\mathbf{P}}\widetilde{\Phi}\widetilde{\mathbf{P}}^T + \widetilde{\mathbf{U}}^2 \tag{3.24}$$

In the formulation above, $\widetilde{\Phi}$ is the inter-factor correlation matrix and is found by

$$\widetilde{\Phi} = \widetilde{\mathbf{T}}\widetilde{\mathbf{T}}^T \tag{3.25}$$

The elements of $\widetilde{\Phi}$ represent the correlations between the rotated factors. As before, $\widetilde{\mathbf{T}}$ is the transformation matrix in loading space that is responsible for rotating the original factors $\widetilde{\mathbf{F}}$ and is determined by the chosen rotation algorithm. $\widetilde{\mathbf{P}}$ is referred to as the "pattern matrix" and is found by utilizing the relationship

$$\widetilde{\mathbf{P}}\widetilde{\Phi} = \widetilde{\mathbf{F}}\widetilde{\mathbf{T}}^T \tag{3.26}$$

Rows of the pattern matrix are essentially regression equations for each measured variable in the study, using the factors themselves as predictor variables. The quantity

$\widetilde{\mathbf{P}}\widetilde{\Phi}$ itself represents what is called the "structure matrix"

$$\widetilde{\mathbf{S}} = \widetilde{\mathbf{P}}\widetilde{\Phi} = \widetilde{\mathbf{F}}\widetilde{\mathbf{T}}^T \tag{3.27}$$

$\widetilde{\mathbf{F}}$ is as it has always been, the matrix of unrotated factor loadings. $\widetilde{\mathbf{S}}$ is thus a matrix of correlations between the items and the rotated factors, whereas $\widetilde{\mathbf{F}}$ was a matrix of correlations between the unrotated items and the factors. Within the factor analytical work done within our study, the structure matrix was utilized. This is due largely to the fact that epistemic beliefs (and hence the factors that define them) may be strongly correlated. The structure matrix, then, would be favored as the correlations between items and factors will not be masked due to any strongly correlated factors. The pattern matrix, however, depends on regression and the impact that one factor may have over another could be diminished during the regression process.

Confirmatory Factor Analysis

Introduction Whereas EFA is said to be a data-driven theory, confirmatory factor analysis (CFA) is said to be a model-driven theory. The idea behind CFA is that you have a theoretical factor structure (usually uncovered with EFA) and you wish to determine the validity of that factor structure. This is done by specifying a factor structure and then using both estimation techniques and new observational data to determine how well the proposed factor structure models the new data. CFA itself is actually a unique case of structural equation modeling (SEM), the covariance structure framework [143].

Mathematical Foundation The general process behind CFA is straightforward, put forth a model and test if that model is adequate. Recall that the idea behind EFA is to reproduce the original correlation matrix using

$$\widetilde{\mathbf{R}} \approx \widetilde{\mathbf{P}}\widetilde{\Phi}\widetilde{\mathbf{P}}^T + \widetilde{\mathbf{U}}^2 \tag{3.28}$$

Within CFA, the factor model is specified by demanding particular elements of $\widetilde{\mathbf{P}}$, $\widetilde{\Phi}$, and $\widetilde{\mathbf{U}}^2$ to either be represented by a variable, by 1, or by 0. As an example, consider the case where a theoretical factor structure is to represent a set of 10 variables ($Q_1 - Q_{10}$) using 3 factors ($F_1 - F_3$). Let us say that items $Q_1 - Q_4$ load (i.e. correlate strongly) with factor 1, $Q_5 - Q_7$ load with factor 2, and $Q_8 - Q_{10}$ load with factor 3. $\widetilde{\mathbf{P}}$ would then be specified by

$$\widetilde{\mathbf{P}} = \begin{array}{c} \\ Q_1 \\ Q_2 \\ Q_3 \\ Q_4 \\ Q_5 \\ Q_6 \\ Q_7 \\ Q_8 \\ Q_9 \\ Q_{10} \end{array} \begin{array}{c} \begin{array}{ccc} F_1 & F_2 & F_3 \end{array} \\ \begin{bmatrix} P_{1,1} & 0 & 0 \\ P_{2,1} & 0 & 0 \\ P_{3,1} & 0 & 0 \\ P_{4,1} & 0 & 0 \\ 0 & P_{5,2} & 0 \\ 0 & P_{6,2} & 0 \\ 0 & P_{7,2} & 0 \\ 0 & 0 & P_{8,3} \\ 0 & 0 & P_{9,3} \\ 0 & 0 & P_{10,3} \end{bmatrix} \end{array}$$

Assuming that oblique rotations of factors are being used (which they are for this work), then the factors should have non-zero correlations and hence $\widetilde{\Phi}$ is represented by

$$\widetilde{\Phi} = \begin{array}{c} \\ F_1 \\ F_2 \\ F_3 \end{array} \begin{array}{c} \begin{array}{ccc} F_1 & F_2 & F_3 \end{array} \\ \begin{bmatrix} 1 & \Phi_{1,2} & \Phi_{1,3} \\ \Phi_{2,1} & 1 & \Phi_{2,3} \\ \Phi_{3,1} & \Phi_{3,2} & 1 \end{bmatrix} \end{array}$$

Lastly, the error is represented by your matrix of uniqueness with

$$\widetilde{\mathbf{U}}^2 = \begin{array}{c c} & \begin{array}{cccccccccc} Q_1 & Q_2 & Q_3 & Q_4 & Q_5 & Q_6 & Q_7 & Q_8 & Q_9 & Q_{10} \end{array} \\ \begin{array}{c} Q_1 \\ Q_2 \\ Q_3 \\ Q_4 \\ Q_5 \\ Q_6 \\ Q_7 \\ Q_8 \\ Q_9 \\ Q_{10} \end{array} & \left[\begin{array}{cccccccccc} h^2_{1,1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & h^2_{2,1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & h^2_{3,1} & 0 & 0 & 0 & 0 & 0 & h^2_{3,9} & 0 \\ 0 & 0 & 0 & h^2_{4,1} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & h^2_{5,1} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & h^2_{6,1} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & h^2_{7,1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & h^2_{8,1} & 0 & 0 \\ 0 & 0 & h^2_{9,3} & 0 & 0 & 0 & 0 & 0 & h^2_{9,1} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & h^2_{10,1} \end{array}\right] \end{array}$$

Here, as mentioned before, the diagonal elements represent the error that is unique to each particular question. To expand upon the example, non-zero elements were included for question/variable 3 with question/variable 9. This is specifying within the model that there is reason to believe these items share some kind of unique error with each other (error that is also not accounted for within the common error).

What remains now is the attempt to model new data with the CFA model presented above. As such, let $\widetilde{\mathbf{R}}_{est}$ represent the corresponding correlation matrix for the above CFA model and $\widetilde{\mathbf{R}}_n$ represent a correlation matrix based on new data (e.g. additional data from 200 more numeric student responses to the 10 questions above). In order to determine how well $\widetilde{\mathbf{R}}_{est}$ compares with $\widetilde{\mathbf{R}}_n$, $\widetilde{\mathbf{R}}_{est}$ is given some initial estimates and the difference between $\widetilde{\mathbf{R}}_n$ and $\widetilde{\mathbf{R}}_{est}$ is found. Iterative techniques are employed until the difference between these two matrices is minimized. This minimization process is done with what is called a "fitting function" (the fitting function is actually what is explicitly being minimized during the iterative process). There exist several fitting functions, for which some common ones include the unweighted least squares, the generalized least squares, and the maximum likelihood. The maximum likelihood function [8] is one of fitting functions used within this work and is defined by

$$F_{ML}(\widetilde{\mathbf{R}}_n, \widetilde{\mathbf{R}}_{est}) = \mathrm{Tr}\,(\widetilde{\mathbf{R}}_n \widetilde{\mathbf{R}}_{est}^{-1}) + \ln\,(\widetilde{\mathbf{R}}_{est}) - \ln\,(\widetilde{\mathbf{R}}_n) - q \tag{3.29}$$

In the above equation, q is a scalar quantity representing the number of free parameters to be estimated (24, in the example) and $\mathrm{Tr}\,(\widetilde{\mathbf{R}}_n \widetilde{\mathbf{R}}_{est}^{-1})$ refers to the trace of $\widetilde{\mathbf{R}}_n \widetilde{\mathbf{R}}_{est}^{-1}$. The other estimator used in this work is the robust weighted least squares (WLSMV), the details of which may be found in its origins paper [150].

A variety of fit statistics are utilized within this work to help quantify how good of a fit the theoretical model provides, however, there are many available in literature. A description of the fit statistics utilized in this study is provided during their presentation. For an excellent account of some common fit statistics, refer to [8, 43].

Correlations

Introduction A correlation coefficient is a numeric value which quantifies the relationship between two variables. This coefficient is typically represented visually by the letter r, and functionally represents how much variation within one variable can be accounted for by considering the variations within another variable. The correlation coefficient utilized with this work is the Pearson product-moment correlation coefficient [7].

Mathematical Foundation

Pearson product-moment correlation coefficient One may quantify how well a particular observed numeric value fits/matches another observed numeric value by taking the difference of the two values. The closer this difference is to zero, the more alike the two values. This basic approach may be expanded to sets of values as well, where each element in a set has a corresponding element in another set. However, a complication occurs in sets of values in that summing the differences of corresponding elements does not yield a reliable measure of fit. This is due to the fact that there will likely be both positive and negative differences that arise within the summation, therefore preventing an accurate measure of how well data match each other. This is remedied by instead summing the square of the difference between corresponding elements (up to n elements), and thus the mathematical formulation of fit is born:

$$Fit = \sum_{i=1}^{n} (modelA_i - modelB_i)^2 \tag{3.30}$$

So the larger this "Fit" the more dissimilar the corresponding values between the two models. This same idea can be applied to a single set of data to help quantify how much the numeric values are spread out from the mean of the data itself. This, referred to as "variance" or "mean squared error", does exactly this but also divides by the number of observations (N) so as to be presented as more of a traditional arithmetic average:

$$Variance = s_x^2 = \frac{1}{N-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 \tag{3.31}$$

The $N - 1$ term is present largely due to the fact that the above variance (s^2) is taken to be a sample variance and not a population variance, and hence $N - 1$ is representing the degrees of freedom within the sample. Combining the above ideas of "Fit" and variance, a measure of how well two data sets with means $\bar{x}$ and $\bar{y}$ vary with respect to each other is given by:

$$Covariance = Cov_{x,y} = \frac{1}{N-1}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) \tag{3.32}$$

This measure, called the "covariance", is one way in which to quantify how two sets of data may share a relationship. As the formula would indicate, when there exist (non-zero) variations in x values and y values, their product is non-zero. The greater and/or more frequent the simultaneous variations in corresponding data elements, the larger the covariance. Thus, by accounting for when data simultaneously deviate from their mean value, we have a way by which to quantify the relationship between two variables. A problem with covariance is its dependency on measurement scale for making comparisons of data, e.g. a covariance of 5 means something different when discussing either miles or kilometers. To resolve issues of measurement error, the covariance between two variables is standardized by now dividing the covariance by the product of the square root of the variance for each set of data:

$$r = \frac{Cov_{x,y}}{s_x s_y} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{3.33}$$

Above is the equation for the Pearson product-moment correlation coefficient, r, a number restricted to values between and including -1 and 1. A value of $r = 1$ would indicate that for every deviation from the mean for one variable measurement (x_i, say) the corresponding deviation from the mean for another variable measurement (y_i, say) is of the exact same value. A correlation coefficient of $r = -1$ would represent much the same, except that the deviation from the mean in each measurement would be of the same magnitude but opposite direction. The Pearson correlation coefficient has the additional benefit in that r^2 may be interpreted as the percentage of variation in one variable (y, say) when the other (x, say) is used as a predictor.

Partial Correlations A partial correlation is much like Pearson correlation in that they both quantify a relationship about the variables involved and are restricted to values between and including -1 and 1. A basic conceptual model of partial correlations may be thought of using three variables x, y, and z. We may ask the question "what is the correlation between dependent variable y and independent variable x?" For which we know that the square of the outcome, $r^2_{y,x}$, would represent the percentage of variation in y when x is used as a predictor. However, if two (or more) predictor variables such as x and z are known to effect y, and x and z share a non-zero correlation, the above may no longer hold true. That is, some of the variation in y may actually be due to z by proxy of variable x. A partial correlation (ρ) seeks to quantify the relationship between two variables when the effects of any other variables are partialed out. In the case of three variables, such as above, the percentage of variation in y when x is used as a predictor and the effects of z are partialed out is $\rho^2_{x,y|z}$, with:

$$\rho_{x,y|z} = \frac{r_{y,x} - (r_{y,z})(r_{x,z})}{\sqrt{1 - r^2_{y,z}}\sqrt{1 - r^2_{x,z}}} \tag{3.34}$$

In the case of more than three variables the mathematics gets slightly more complex, however the interpretation remains the same [41].

Hypothesis Testing In determining whether a Pearson correlation is significantly different from zero, a Student's t sampling distribution may be assumed with $\nu = n-2$ degrees of freedom, and a test statistic of:

$$t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}} \tag{3.35}$$

Alternatively, we could make use of the Wald test statistic [206] for a single parameter sampled from a normal distribution, for which:

$$W = \frac{\hat{\theta} - \theta_o}{SE(\hat{\theta})} \tag{3.36}$$

$\hat{\theta}$ is a maximum likelihood estimate of the sampling distribution parameter to be tested and θ_o is the proposed value of the same parameter from the sampling population. In more familiar terminology, a z-score may be utilized:

$$z = \frac{\bar{x} - \mu}{SE(x)} \tag{3.37}$$

Which is essentially a Wald test statistic with $\bar{x}$ (the arithmetic mean) being the maximum likelihood estimate of x, a value drawn randomly from a normal distribution. The problem with correlations, however, is that the sampling distribution of r

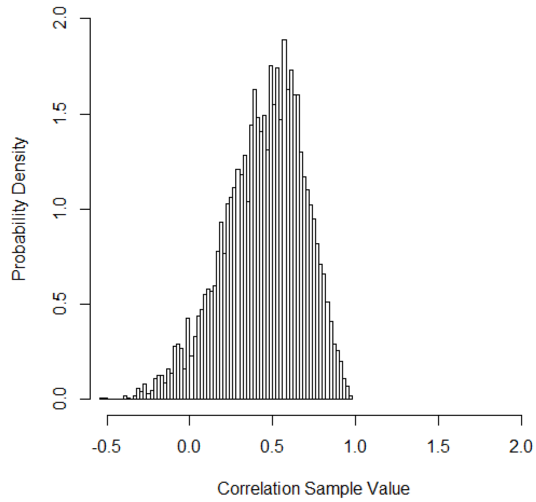

Figure 3.1: Sampling distribution of Pearson r correlation coefficients from a sample in which $r = 0.445$. Summing all bins has value unity.

has a tendency to be skewed (e.g. Figure 3.1). To deal with distributions which are skewed, a Fisher transformation [78] is employed:

$$r_z = \frac{1}{2} \ln \left(\frac{1+r}{1-r} \right) = tanh^{-1}(r) \tag{3.38}$$

This transformation will adjust r such that the sampling distribution now be approximately normal (e.g. Figure 3.2). Thus an observed correlation value r may be transformed into r_z where we may obtain a z-statistic via:

$$z = \frac{r_z - \mu}{SE(r_z)} \tag{3.39}$$

For this sampling distribution it was found that $SE(r_z) = \frac{1}{\sqrt{N-3}}$ [79], where N is the number of samples from which r_z was determined. As the tests done within this work are to determine if these correlations are significantly different from zero, the

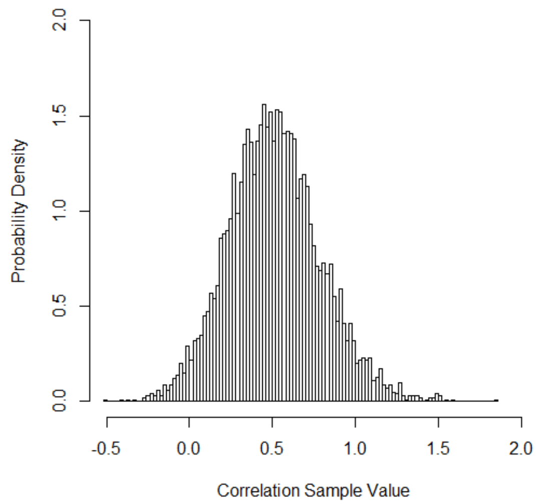

Figure 3.2: Sampling distribution of Fisher-transformed correlations, r_z, from an original sample in which $r = 0.445$. Summing all bins has value unity.

null hypothesis would have $\mu = 0$. The z-statistic is then:

$$z = \frac{r_z}{\frac{1}{\sqrt{N-3}}} \tag{3.40}$$

The above z-statistic is intended for use with Pearson correlation coefficients. Should partial correlation coefficients be tested for significance it is recommended that N become $n - s$, where s refers to the number of variables partialed out [80].

Non-parametric Methods

Introduction Many of the common statistical hypothesis tests rely on the primary assumption that the data sampled are from a normal distribution. When this cannot be upheld, other methods are sought which most typically involve ranking the data numerically. By then conducting an analysis on the rankings instead of the raw data, issues such as distribution skew and outliers are eliminated.

Wilcoxon Tests

Wilcoxon signed-rank test As the name suggests, the Wilcoxon signed-rank test [210] utilizes rankings of data in order to conduct tests of significance. This test is designed to compare whether two related samples of data (e.g. matched pretest and post-test scores of students) are drawn from the same distribution. The null hypothesis for this test would be that the difference in the scores of the related samples of data are centered about zero with a symmetric distribution. The general idea is that a difference in the related scores are acquired, then the absolute value of these scores are given a numeric rank from lowest to highest. The process begins with finding D_i:

$$D_i = |x_i - y_i| \tag{3.41}$$

Where x_i and y_i represent the related data samples ($i = 1, 2, 3, \ldots N$, N is the total number of related data samples), and for which values of $D_i = 0$ are excluded. The values of D_i are then ranked from lowest to highest, with the lowest value being assigned a 1 and incrementing by 1 (if possible, see following) for each new rank. Values of D_i which are the same are given a rank which is equal to the mean of the range over which they span. For example, the values $D_i = \{2, 6, 6, 6, 6, 10, 13\}$ would be given ranks of $Z_i = \{1, 3.5, 3.5, 3.5, 3.5, 6, 7\}$. The Wilcoxon signed-rank test statistic is:

$$W = \sum_{i=1}^{n} Z_i R_i \tag{3.42}$$

In the equation above, n is the number of values of D_i which were non-zero, Z_i is the rank assigned to D_i, and R_i is a value which is either -1 (if $x_i - y_i < 0$) or 1 (if $x_i - y_i > 0$). It is of note that some treatments of R_i are such than it is either 0 (if $x_i - y_i < 0$) or 1 (if $x_i - y_i > 0$). As n becomes large (as in this study) the sampling distribution of W approaches that of a normal distribution with an expected value of 0 (or $\frac{n(n+1)}{4}$, if non-zero and non-negative ranks are used) and a variance of:

$$\hat{\sigma} = \sqrt{\frac{n(n+1)(2n+1)}{24}} \tag{3.43}$$

Recall from earlier that there exists an advantage for sampling from a normal distribution (Wald statistic), in that one can define a z-score. For samples of W, it follows:

$$z = \frac{W}{\hat{\sigma}} \tag{3.44}$$

or (if non-zero and non-negative ranks are used):

$$z = \frac{W - \frac{n(n+1)}{4}}{\hat{\sigma}} \tag{3.45}$$

MannWhitney U test The Mann-Whitney U test and the Wilcoxon rank-sum test are two equivalent statistical tests developed independently in 1947 and 1945, respectively [137, 210]. Like the Wilcoxon signed-rank test, the Mann-Whitney U test makes use of rankings of the data. Both these tests are akin to an independent samples t-test, from which raw data come from a normal distribution, and employs a Student's t distribution. The null hypothesis being tested within this study is that the distribution of both groups in question is identical.

For brevity only the Mann-Whitney U method will be discussed, which begins by assigning ranks to the data from pooled data. So, for example if group A had values $A = \{3, 6, 8, 11\}$ and group B had values $B = \{4, 6, 12\}$ then the rankings for group A (R_A) and group B (R_B) would be $R_A = \{1, 3.5, 5, 6\}$ and $R_B = \{2, 3.5, 7\}$, respectively. As seen with rankings in the Wilcoxon signed-rank test, data are ranked from lowest to highest, with the lowest value being assigned a 1 and incrementing by 1 (if possible, see following) for each new rank. Data which are numerically identical are given a rank which is equal to the mean of the range over which they span (as demonstrated in the aforementioned example). Continuing with the example, the U statistic for each data set is then defined as:

$$U_A = R_A - \frac{n_A(n_A+1)}{2} \qquad U_B = R_B - \frac{n_B(n_B+1)}{2} \tag{3.46}$$

Where n_A and n_B are the sample sizes for groups A and B, respectively. The total sum of all the ranks is $\frac{N(N+1)}{2}$, with $N = n_A + n_B$. Conceptually, the value U_A (and U_B) represent the observed rank-sum minus the theoretical rank-sum (if the data were sufficiently far apart such that the ranks of each group were independent subsets e.g. $R_A = \{1, 2, 3, 4\}$ and $R_B = \{5, 6, 7\}$). The smaller of the U statistics is considered the test statistic. For large samples ($N > 80$ or each $n_i > 30$ [44, 56]), U approximately follows a normal distribution with mean $U_o = \frac{n_A n_B}{2}$ and standard deviation:

$$\hat{\sigma} = \sqrt{\frac{n_A n_B (N+1)}{12}} \tag{3.47}$$

Should there be ties which occur in ranking, the adjusted standard deviation is:

$$\hat{\sigma} = \sqrt{\frac{n_A n_B}{N(N-1)} \left(\frac{N^3 - N}{12 - \sum_{i=1}^{g} \frac{t_i^3 - t_i}{12}} \right)} \tag{3.48}$$

Where t_i represents the number of occurrences in which a subject shares a specific rank (symbolized by i) and g is total number of groups of ties. Of course, as U now approximately represents a normal distribution it follows that one can define a z-statistic as:

$$z = \frac{U - U_o}{\hat{\sigma}} \tag{3.49}$$

Should the critical value imposed on z result in rejecting the null hypothesis, it is sufficient to say that the populations are different in some way (center, spread, or shape). When the distributions of raw data are similar, then rejecting the null hypothesis is interpreted as one group having a significantly different median than the other.

Chi-Squared Tests Chi-squared tests are statistical procedures designed to determine if there exists a significant difference between the observed frequencies/counts of occurrence within a set of data compared to what is expected (as defined by the researcher). The Pearson chi-squared test [162] is designed to determine if the categorical variables within a population are independent, under the assumption that the distribution being drawn from is a chi-squared distribution.

Chi-squared test of homogeneity The chi-squared test of homogeneity is the only chi-square test utilized within this work. This test is applied to a single categorical variable (which could have multiple levels) sampled from two or more populations, with the null hypothesis being that the proportion of observations for each variable is identical across populations. Within this study, for example, the categorical variable was a set of response types (0-unsophisticated, 1-neutral/middling, 2-sophisticated) to a given EBAPS question across several populations (baseline data and modified data). The chi-squared homogeneity test statistic for samples drawn from a chi-squared distribution is:

$$\chi^2 = \sum_{i=1}^{n} \frac{(\text{observed- expected})^2}{\text{expected}} = \sum_{i=1}^{c}\sum_{j=1}^{p} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \tag{3.50}$$

In the above equation the sum is over all c levels of a category and every count (up to the pth count) for a population within that category. The expected count is given by:

$$E_{ij} = \frac{N_i N_j}{N} \tag{3.51}$$

Where N_i is the total number of observations from category level i, N_j is the total number of observations from population j, and N is the total sample size. One could think of the expected count as being the number of observations from a category level multiplied by the probability of belonging to a given population (given by N_j/N). As an example, if you consider flipping a coin 20 times, the expected count of observing heads would be $20 * (\frac{10}{20}) = 10$. The observed count is, of course, the number of occurrences that the researcher observes from population j in category level i. The form of the chi-square distribution that the test statistic χ^2 is utilized for depends upon the number of degrees of freedom in the problem, given by:

$$df = (c-1)(p-1) \tag{3.52}$$

For which c is the number of levels for the categorical variable and p is the number of populations.

Analysis of Variance (ANOVA)

Introduction An analysis of variance (ANOVA) is a statistical procedure designed to test for the existence of at least one significant difference between the means of two or more groups [79, 81]. It follows that the null hypothesis for such a procedure would be that there exist no significant differences between the group means in question.

Mathematical Foundation The test itself is straightforward, and begins with calculating the total sum of squares for all the groups:

$$SS_t = \sum_{i=1}^{g} \sum_{j=1}^{n_g} (x_{ij} - \bar{x}_{grand})^2 \tag{3.53}$$

SS_t represents a summation over all of the differences between the jth value of the ith group and the grand mean $\bar{x}$, quantifying how much the data is spread about the grand mean. The grand mean itself is simply the mean of all the data involved in the analysis. Much like SS_t was done with respect to the grand mean, the model sum of squares is done with respect to both the group and grand means:

$$SS_m = \sum_{i=1}^{g} n_i(\bar{x}_i - \bar{x}_{grand})^2 \tag{3.54}$$

With n_i representing the number of values within the ith group. From inspection of SS_m it is clear that should all groups have a mean that is identical to the grand mean, then then SS_m would be zero (the group means perfectly "predict" the grand mean). Should these differences be non-zero, then the value $SS_t - SS_m$ would represent the amount of variation not explained by the model (the group means) and may otherwise be calculated with:

$$SS_e = \sum_{i=1}^{g} \sum_{j=1}^{n_g} (x_{ij} - \bar{x}_i)^2 \tag{3.55}$$

SS_e (error sum of squares) is similar to SS_t with the exception that the differences are now taken with respect to the group mean/s. It would follow that one could acquire a signal-to-noise ratio dividing SS_m by SS_e. As each of these values is determined with a different number of data, they are standardized by dividing by their respective degrees of freedom. That is:

$$MS_m = \frac{SS_m}{df_m} \qquad MS_e = \frac{SS_e}{df_e} \tag{3.56}$$

MS is referred to as the "mean squares". The degrees of freedom for MS_m is $df_m = g - 1$, where g is the number of groups within the study. The degrees of freedom for MS_e is $df_e = n - g$, where n is the total sample size for the study. It follows that a signal-to-noise ratio , F, would be:

$$F = \frac{SS_m}{SS_r} \tag{3.57}$$

This value, F, is actually the test statistic for a procedure known as "Fisher's F-test", a statistic treated as having been drawn from a F-distribution.

Recall that an ANOVA will only reveal if a significant difference exists between the group means, it will not reveal which group/s are creating this significance. In order to determine which means are significantly different from which other means, a post-hoc test is conducted. Two of the prominent post-hoc tests conducted for ANOVA follow-up are the Tukey and Duncan tests [61, 203]. Within this work a Duncan post-hoc test (Duncan's multiple range test) was conducted, as it is a more powerful test (more likely to detect significance when it exists) but at the risk of an increased likelihood that significance be detected when it does not exist [28, 67]. To guard against this aggressive choice, a conservative Bayesian approach was also utilized, known as the Waller-Duncan test. Details regarding the algorithms for each of these approaches may be found in [61] for the Duncan post-hoc test and [207] for the Waller-Duncan post-hoc test.

RESULTS

Exploratory Factor Analysis Within SPSS

This section, titled "Exploratory Factor Analysis Within SPSS" is an adaptation of previously published material. The following is functionally the same material as that which was originally submitted and accepted to the peer-reviewed journal *Physical Review Physics Education Research* (volume 14, issue 1, pages 010135, co-author/s: Shannon D. Willoughby), with the exception of some basic formatting modifications [116].

Introduction

This paper is a continued exploration of student epistemological beliefs within introductory astronomy as presented in previous literature [116, 212]. The goal is to identify and analyze response patterns of students in an introductory astronomy class using exploratory factor analysis on data from the epistemological beliefs assessment for physical science (EBAPS). We believe the findings will help further understanding of student epistemologies in introductory astronomy, an area where literature is lacking. This analysis also provides some insight into the EBAPS instrument itself. Fundamentally, we seek to answer the following research question: What factor structure is present for astronomy students who take the EBAPS and what do these factors represent?

The EBAPS suffers from a distinct lack of validity studies within literature, as others have also noted, despite its prominence as one of the more well-known epistemic instruments within physics education research [10,59,60,149,157]. It is important that instruments as influential as the EBAPS experience multiple validation studies so as to further establish the legitimacy of what the instrument claims to assess [105,124,164].

We do not claim that this work is a complete and thorough validation study of the EBAPS (validation of this assessment is expanded upon in the "EBAPS" section of the following page), but rather we seek to gain insight into how students are responding to the assessment by exploring the constructs present within the instrument and how they relate to the constructs as proposed by the authors of the assessment.

The analyses done herein, which utilize five years of posttest data from an introductory astronomy course, will begin to explore the psychometric properties of the EBAPS. The factor structure revealed by the exploratory factor analysis done will be further refined using written student responses of items belonging to these factors. Last, the implications of these findings will be discussed as will future work regarding continued validation work and the EBAPS.

Methods

EBAPS The epistemological beliefs assessment for physical science (EBAPS) is a 30-item forced-choice instrument designed to assess students epistemologies within the physical sciences. The EBAPS was developed by a research group (Andrew Elby, John Frederiksen, Christina Schwarz, and Barbara White) from University of California, Berkeley. The instrument is designed to focus attention on personal epistemologies as opposed to expectations [69]. The authors claim the EBAPS assesses five nonorthogonal dimensions of epistemological beliefs: Structure of scientific knowledge, nature of knowing and learning, real-life applicability, evolving knowledge, and source of ability to learn. Precise definitions of these axes have been provided in Table A.1 in Appendix A. Additional information regarding the EBAPS may be found on the host website [68]. Each item on the EBAPS is scored on a scale from 0 to 4, where 0 represents an unsophisticated view and 4 represents a more sophisticated or more expertlike view. Questions are mapped onto one of the five axes defined by the authors of the instrument. Axis scores represent an average of all item scores belonging to an axis; which items correspond to a particular axis may also be seen in Table A.1.

The authors of the EBAPS explicitly state that they "[...] dont want to assume that each subscale corresponds to a stable, consistent belief (or set of beliefs)" and provide compelling reasoning as to why they take this stance [69]. In essence, the authors of the EBAPS follow a cognitive theory which consists of "fine-grained cognitive resources" which are activated upon context, similar to DiSessas p prims [96]. What this means is that items that belong to the same subscale, category, or axis may be answered differently depending on particular contextual cues, but in a general sense should still be placed within a certain subscale. It may be worth wondering why bother promoting axes at all within the instrument. However, experts may still provide a general category for which these items should belong; this categorization along axes is an approach many epistemic instruments within physics education research follow. Whether the constructs (beliefs) that are found in this study represent a contextual activation of local cognitive resources or they represent fundamental beliefs with varying levels of sophistication, they may be used to help better understand and address student epistemologies within introductory astronomy.

Context and Participants The EBAPS was given as a pre-test and post-test in a large enrollment introductory astronomy course at a medium-sized, midwestern, land-grant institution. The course had active learning elements and typically consisted of a 45 minute lecture with a 30 minute group activity. Lecture itself made frequent use of iClickers, writing activities, and student-student interactions while group activities made use of the tutorials for introductory astronomy to supplement lecture [168]. The data in this study were taken from fall of 2012 through spring of 2017 across two identical sections of the course, 2334 students in all. Although precise information of classroom demographics over this time could not be acquired, in fall of 2017 the university as a whole had an ethnic distribution which was 83% Caucasian, 4% Hispanic, 2% American Indian, and 1% Asian. The remaining 10% were African American, Pacific Islander, or a combination of multiple ethnicities. Roughly 47% of students were female and 53% were male. The instructor notes that each section consisted of nearly 200 students a semester and was predominantly Caucasian with an approximately equal distribution of males and females. The same instructor (one of the authors) taught the course across both sections and all five years of the study.

Exploratory Factor Analysis Factor analysis is a statistical method that transforms a set of variables into factors whose items contain some type of underlying similarity. In general, there are three reasons for conducting an exploratory factor analysis (EFA). The first is to discover the underlying structure responsible for guiding student responses. This is what the current paper will focus on, as the underlying epistemic belief structures influencing student responses to the EBAPS are sought. The second is to reduce a set of data into simplified components. For instance, if a set of variables are being utilized to measure a trait such as intelligence, there could arise a scenario in which several of these variables have strong linear relationships with one another. Here the second method, typically known as principal components analysis, will reduce these variables of high multicollinearity into a single predictor. Last, EFA may be used in the creation of an instrument, assisting the researcher in determining whether the variables put forth are measuring the latent variable they intend to measure. It is not unusual for an EFA to be followed up with a confirmatory factor analysis (CFA). Briefly, a CFA is a type of factor analysis that measures the consistency of any factors that exist within a model proposed by the researchers. A CFA may be used on the same data as an EFA to help refine items and/or item relationships. Alternatively, a CFA may be used on a separate set of data as an EFA to determine how well the proposed model fits with alternate data.

The authors of the EBAPS have attributed values (between 0 and 4) to each question within the instrument. EFA traditionally utilizes the Pearson correlation matrix, which is based on the assumption that the data exhibit an equal interval scale with linear relationships between items. However, many items on the instrument have a Likert-style response and as such cannot necessarily be assumed to exhibit these traits [200]. Violations of this assumption (i.e., Likert-style data) could potentially result in Pearson correlations that underestimate the strength of correlations between items thereby leading to improper factors and/or factor loadings [21,155]. Alternative options could include the use of polychoric correlations with EFA. Polychoric correlations are specifically designed to estimate correlations for ordinal data from a bivariate normally distributed population. In the end, it was decided to implement traditional EFA but with the principal axis factoring extraction methodology, which makes no assumptions of the underlying distribution of the data [75]. This same approach to EFA has been utilized on other epistemic instruments within the sciences as well [58].

Ideally, if proper factors underlying student epistemic beliefs in astronomy can be identified, then future work may focus on addressing these traits within the classroom, either for improvement or refinement. The EBAPS itself may benefit from this work as well, given that more data regarding response behavior to the instrument will be acquired.

Data Analysis The following section is composed of two parts, first determining the number of factors to retain, and then assessing the stability of the factors. All analyses within were done utilizing IBM SPSS software. This provided a well-vetted template from which to conduct work and allowed for easy partitioning and thorough exploration of the data.

Number of factors to be retained Parallel analysis and scree plots were generated to assist in determining the number of possible factors involved. Also utilized was an EFA with the Kaiser criterion method, which states that factors having eigenvalues greater than or equal to 1 are to be retained [118]. We proceed by discussing the results of all three methods. The methods below make use of the post-test data of students from section one, which had 1258 students across five years. Section two will later be utilized in future work with confirmatory factor analysis.

A scree plot takes the factors from a correlation matrix of assessment items and plots them against their corresponding eigenvalues. In a scree plot it is typical that the eigenvalues associated with most factors tend to lie along a single line. To determine which factors to retain, look for when the slope of this line makes a distinct change and keep the factors whose eigenvalues rise above that line. The scree plot generated from the 30 EBAPS items, seen in Fig. 1, indicates three prominent factors with upwards of six possible factors (seven if the inflection point is included) [31].

Parallel analysis involves random generation of pseudoclasses (based off of raw data input) from which a correlation matrix is created and its factor eigenvalues calculated. These eigenvalues are thus values arising from random noise within the raw data. As such, only the eigenvalues from the factor analysis that are larger than these noise-based eigenvalues are not likely to be artificial constructs created from deviations in data. The parallel analysis was done using an add-on to SPSS that can be obtained online [3, 153]. Eigenvalues generated in the analysis were done with permutations of the raw 1258-student data set using Castellan's algorithm [30]. The results of the parallel analysis were similar to that of the scree plot, revealing three

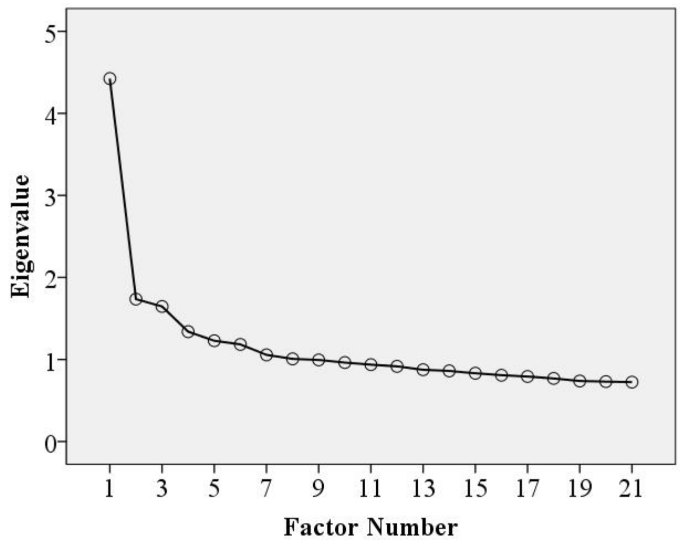

Figure 4.1: Scree plot for the 30 EBAPS items (post-test). Factors after 21 were truncated.

to five prominent factors, but up to seven possible factors overall.

In the final attempt to determine the number of axes to retain, an exploratory factor analysis using the Kaiser criterion was performed with all students from section one and using all thirty items of the EBAPS. In doing so, this analysis did not restrict the number of factors that could be present. This method also did not rely purely on the number of eigenvalues greater than unity to determine factors, but instead looked at how many eigenvalues were over unity after other filtering criteria (detailed later) were applied. The EFA implemented the direct oblimin oblique rotation method. An oblique rotation was utilized as we expect some general correlation of epistemological constructs [75]. Other often-used oblique rotations include the quartermin and promax methods, although there is reason to believe that all these oblique rotations lead to a similar outcome [75].

Principal axis factoring was utilized instead of principal components analysis (PCA) as we sought to find underlying relationships between items and avoid distributional assumptions, as opposed to a PCA, which is utilized to reduce data into simplified minimal components [25, 47, 51, 196]. In conducting the EFA the determinant of the correlation matrix for this data was nonzero (determinant = 0.026), suggesting that linear combinations of data (i.e., factors) are possible. Furthermore, Bartletts test of sphericity indicated that this determinant is indeed significantly different from zero ($p < 0.001$). The Kaiser-Meyer-Olkin (KMO) test revealed that the degree of common variance among the items was acceptable as Meritorious (KMO = 0.846) [74]. Last, the diagonal elements of the anti-image correlation matrix were all above 0.50, a sign that these items may be suitably explained by underlying factors [33]. Some issues did arise when parsing the correlation matrix, however. Ideally an item should have at least one correlation exceeding 0.30, however, several items within the data did not reach this cutoff [196].

Upon analyzing the correlation matrix, we chose to remove items which did not have a correlation coefficient greater than 0.2. This value was chosen instead of 0.3 because most items either exhibited at least one correlation greater than 0.2 or were notably far beneath this measure. The items 13, 18, 19, 20, 23, 24, 26, and 29 were removed from the factor analysis during this process, and the analysis was redone. Ideally all items in a survey can be kept, however, it is common for some items to simply not group strongly enough to indicate the presence of a prominent latent variable. Item 2 was removed in this second iteration when it was found that it had a loading less than 0.32 [196]. Again the analysis was redone and at this point the communalities of the items were inspected. Items whose communality was less than 0.20 were removed one at a time, with the lowest value being removed first, redoing the factor analysis after each item was removed [35]. This led to the removal of items 3, 7, 15, 17, and 27. The pattern matrix for the final iteration of this exploratory factor analysis is displayed in Table 4.1 ; also provided within Table

4.2 are the corresponding factor correlations for the analysis. As can be seen, there are four primary factors left and none of the fifteen items exhibited a crossloading greater than 0.40 [88]. All but one factor retained at least three items which is a recommended criteria for factor stability [91]. This final factoring also upheld all basic assumptions, namely, a KMO = 0.845, and Bartletts test of sphericity having $p < 0.001$.

Given the findings for the Kaiser criterion factor analysis, scree plots, and parallel analysis, it has been decided that both a three-factor model and a five-factor model are to be presented. We neglect formally discussing this potential four-factor Kaiser criteria-based model because the five-factor model does a satisfactory job of retaining and yet adding on to the information provided by the four-factor model. A six-factor model was also considered, however, some compromises to commonality criteria as well as crossloading issues with item 12 led to the dismissal of the model. Beyond that, models of seven and eight factors were explored but it was found that loading and crossloading issues became too frequent to retain viable independent factors. It is possible that future work will reconsider a four-factor and six-factor model with confirmatory factor analysis. Future CFA work will also consider items that were removed from their original factors due to the strict filtering criteria applied here.

Factor Stability Having decided the number of factors to retain, an EFA comparison was conducted for both a three-factor model and a five-factor model. To test item stability, students from the above section one data were split into two roughly equal groups for both the three-factor model and five-factor model; the groups were then compared to one another [156]. The split itself was based alphabetically by last name, as there is no reason to believe that a strong correlation exists between last name and epistemic beliefs within this demographic. The number of students from

Table 4.1: Factor Loadings from EFA Utilizing the Kaiser Criterion. Loadings less than 0.3 have been suppressed. EFA involved all section 1 students. These factors represent factors whose eigenvalues are over unity after other filtering criteria were applied to the original 30 items.

Items	Factor 1	Factor 2	Factor 3	Factor 4
14	0.583			
08	0.556			
06	0.467			
10	0.442			
04	0.430			
22		0.719		
21		0.603		
25		0.415		
30			0.593	
28			0.417	
16				0.558
05				0.535
11				0.513
12				0.437
09				0.353

Table 4.2: Factor Correlation Matrix Results for EFA Utilizing the Kaiser Criterion. EFA involved all section 1 students. These factors represent factors whose eigenvalues are over unity after other filtering criteria were applied to the original 30 items.

Factors	1	2	3	4
1	1.000	—	—	—
2	0.310	1.000	—	—
3	0.451	0.347	1.000	—
4	0.619	0.206	0.313	1.000

each group is not consistent, as items between groups and models are not similar and SPSS does not consider students with incomplete data for an item. Roughly, the first group (group A) contained around 580 unique students and the second group (group B) contained around 630 unique students. The approach in factor simplification within these analyses utilized a similar methodology as was done for the Kaiser criteria factor analysis above except the number of factors were simply restricted in SPSS to be 3 and 5, accordingly. It is worth noting that some items removed via this process may again be considered during CFA model refinement. For now, we are utilizing this process to provide the most mathematically sensible model possible with EFA.

Three-factor model. The initial results for the three-factor models of both groups A and B are shown in Tables 4.3 and 4.4, respectively. A direct comparison between these models reveals similar factors with mostly consistent item loadings. As we seek a dependable theoretical model, dissimilar items were removed. This resulted in items 2 and 4 being removed from factor 1 in group A while items 28 and 30 were removed from factor 3 in group B. Similarly, item 3 was also removed from factor 3 in group A.

An EFA was redone for each group utilizing only the remaining common items, the results of which are shown in Table A.2 (group A) and Table A.3 (group B) within Appendix A. We thus propose a theoretical three-factor model shown in Table 4.5.

*Five-factor model.*Table 4.6 (group A) and Table 4.7 (group B) reveal the results for each group within the five-factor fitting. As can be seen, there are modest discrepancies between both results. However, prominent factors should be similar between the three-factor theoretical model and these five-factor theoretical models. As a result, each group model was reworked. Group A had item 1, 2, and 3 removed as these items were not present in the theoretical three-factor model, nor are items 1 and 2 present in the five-factor model of group B. Item 16 was added to group A as were items 28 and 30 due to their presence in both group B and in their previous association with the three-factor model of group A (they were initially removed due to weak correlations, $r < 0.25$). For similar reasons, item 10 was added to group B while item 7 was removed. Although these were not minor changes, results indicate relative agreement between the refined 5-factor models of group A (Table A.4) and group B (Table A.5) as seen in Appendix A. Utilizing these findings, a five-factor theoretical model is hence put forth in Table 4.8. Despite inconsistencies between the loading of items 4 and 9, as well as the concerning presence of factors with fewer than three items, we will discuss shortly how student responses have provided insight as to why items 4 and 9 may represent some form of underlying belief.

It is important to consider all models presented within this section as future CFA work may reveal that one particular model is more robust when applied to any alternative student data. Ideally, the five-factor model would have adequate fit statistics in CFA and be comparable to the three and four-factor models. Should that be the scenario, the five-factor model will be favored as it details a larger swath of the collective students epistemic belief structure.

Table 4.3: Initial 3-Factor Loadings for Group A. Loadings less than 0.25 have been suppressed. EFA involved approximately half of the section 1 students. These loadings represent results after a filtering of the original 30 items.

Items	Factor 1	Factor 2	Factor 3
08	0.780		
14	0.564		
02	-0.466		
06	0.465		
10	0.423		
04	0.390		
22		0.737	
21		0.617	
25		0.405	
11			0.556
05			0.539
03			0.477
16	0.262		0.411
12			0.405
09			0.356

Table 4.4: Initial 3-Factor Loadings for Group B. Loadings less than 0.25 have been suppressed. EFA involved approximately half of the section 1 students. These loadings represent results after a filtering of the original 30 items.

Items	Factor 1	Factor 2	Factor 3
16	0.610		
09	0.505		
12	0.487		
05	0.464		
11	0.433		
21		0.642	
22		0.603	
25		0.465	
08	0.258		0.465
30			0.452
28			0.450
14			0.407
10			0.398
06			0.377

Table 4.5: Theoretical 3-Factor Model

Factor 1	Factor 2	Factor 3
Question 06	Question 21	Question 05
Question 08	Question 22	Question 09
Question 10	Question 25	Question 11
Question 14		Question 12
		Question 16

Results

To assist in the interpretation of the proposed factors presented in Tables 4.5 and 4.8, written student responses were utilized. In the latter years of the study students had access to an online homework site called Sapling Learning [1]. As such, EBAPS questions were able to be presented to the students in such a manner that they could type an open-ended response to the question asked. Their response was voluntary and no credit was awarded for participation. From these student responses, information pertaining to all factor analysis questions in both models was able to be obtained (with the exception of question 4). The responses, which typically numbered between 40 and 90 responses per question, were analyzed and partitioned for various patterns. It is advised that the reader utilize the EBAPS alongside the discussions, as there are frequent interpretations of the questions themselves. The discussion will predominantly involve an analysis of the five-factor theoretical model, as it has many similarities with the three-factor model while containing more information. That is, both models have three nearly identical factors which test the same latent variable, but the five-factor model also includes an additional two factors.

Table 4.6: Initial 5-Factor Loadings for Group A. Loadings less than 0.25 have been suppressed. EFA involved approximately half of the section 1 students. These loadings represent results after a filtering of the original 30 items.

Items	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
08	0.721				
02	-0.554				
06	0.471				
14	0.430				
10	0.324				
22		0.743			
21		0.651			
25		0.359			
11			0.615		
05	0.250		0.414		
03			0.386		
01				0.669	
12			0.354	0.354	
09					-0.661
04					-0.389
16			0.272		-0.284

Table 4.7: Initial 5-Factor Loadings for Group B. Loadings less than 0.25 have been suppressed. EFA involved approximately half of the section 1 students. These loadings represent results after a filtering of the original 30 items.

Items	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
16	0.594				
05	0.523				
09	0.471		0.271		
11	0.464				
12	0.445				
03	0.279				
22		0.735			
21		0.573			
25		0.461			
04			0.542		
30				0.590	
28				0.367	
06					0.539
07					0.415
08					0.404
14					0.376
10					0.336

Table 4.8: Theoretical 5-Factor Model

Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Question 06	Question 21	Question 05	Question 28	Question 04
Question 08	Question 22	Question 11	Question 30	Question 09
Question 10	Question 25	Question 12		
Question 14		Question 16		

Theoretical Five-Factor Model

1. Factor 1: Structure of science

The first factor in the theoretical five-factor model includes EBAPS questions 6, 8, 10, and 14. The general theme seen in student responses here may be classified as *network of science* or *structure of science*. The majority of written responses to question 14 involve the idea that science permeates everyday life and is thus a necessary field in which politicians should be familiar. Question 8 written responses discuss the belief that theories act as a webbing which connects and accounts for information. These theories are thus vital for the accumulation of knowledge within science. Written responses for question 6 tend to have students focus on the scientific community and the idea that science is a reliable, rigorous process or entity. Last, question 10 written responses convey a type of trust-in-science mindset, exemplified by many students making the conclusion that if a scientific theory or principle does not make sense, then the student must be missing something. Again, basing the idea that the theory must be capable of being understood because science created it and science is a rigorous and reliable process.

Note that item 10 is likely grouping to this factor because of the latter portion of the question, which states "[. . .] not everything in science is supposed to make sense." Question 10 was designed to address the "structure of scientific knowledge" (see Table A.1) but some responses did indicate a metacognitive ("nature of knowing and learning") aspect which focused on the "[. . .] you have to accept it and move on [. . .]" question segment.

Overall, this factor seems to be testing student views of the philosophical, psychological, and occasionally sociological aspects of the nature of science (NOS). Aspects of NOS such as "science is an attempt to explain phenomena," "scientists work collaboratively," "science aims to be consistent," and "science relies on skepticism" were often seen within the written student responses [142]. In conclusion, it may be best to categorize this factor as *structure of science*, given that student responses largely display a belief of science as a network of people, processes, or information involved in understanding everyday life. This is distinct from structure of scientific knowledge but not exclusively so. Written responses often involved discussions of the nature of science, which may then manifest in their views regarding the Structure of scientific knowledge. Epistemologically, this factor could be probing student views with respect to the philosophical and psychological aspects of the nature of science.

2. Factor 2: Innate ability vs hard work

Questions 21, 22, and 25 make up factor two within the five-factor model. These questions are quite likely probing the extent to which students attribute hard work or natural ability to achievement within science, and life in general. Although no explicit responses were requested for questions 21 and 22, students were asked to define natural ability and hard work. Written response analysis revealed students categorized hard work as dedication to learning the material, perseverance, and/or working to fully understand the material. In summary, one may say students see hard work as dedication and perseverance toward understanding the material. Natural ability, however, was less agreeable in its categorization. The first, and most obvious, response was that natural ability is a trait governing scientific proficiency and that you are born with it. Other students did not clarify if natural ability was truly nature or nurture, but did associate it with an ability to learn the material while having put in considerably less work than their peers. Several of these responses specifically referred to being able to learn the material quickly. Last, a non-negligible number of students referred to natural ability as being something that a student has acquired and brought into the classroom; these students associated natural ability with either prior knowledge, work ethic, and/or interest in the subject. Written responses to question 25 reinforced these findings and further displayed the dichotomy of hard work versus natural ability.

Although this factor seems best suited to be labeled *innate ability vs hard work*, it does seem to still partially align with "source of ability to learn" which the EBAPS associates with questions 22 and 25. Interestingly, there were also traces of Schommer's quick learning present within these questions (via the student written responses) as well, specifically in how students viewed natural ability [183,184]. That is, typical student responses associated natural ability with being able to either learn the material quickly or on their first attempt [183,184]. Epistemologically, this factor appears to pit the influence of hard work against that of natural ability when assessing the role that they play in student views of success.

3. Factor 3: Source of ability to learn

Factor three in the five-factor theoretical model includes EBAPS items 5, 11, 12, and 16. According to the EBAPS (Table A.1), the items in this factor are either nature of knowing and learning (11 and 12) or source of ability to learn (5 and 16). Upon analyzing student responses, the one clear trend for all these items is the belief that science can be learned.

Written response data for question 5 focuses on the belief that there are a variety of ways one can learn material, naming strategies such as memorization, incremental studying, and group studying. The same overarching belief, that science can be learned, was present for question 11 in the form of content familiarity. In particular, familiarity with test material as developed via time devoted to studying. Epistemologically, question 12 responses focused on two segments. Based on written response data students likely read item 12 as either "When learning science, it is possible for people to understand the material" or "When learning science, it is beneficial to relate the material to your own ideas." The latter here would test nature of knowing and learning, whereas the former is more closely related to source of ability to learn. This is further made obvious as question 1 occasionally grouped with question 12 in the many approaches or iterations to factor analysis that were utilized (as seen outside of what has been presented within this paper).

Question 1 clearly tests one prominent aspect of epistemic beliefs (as seen with written response data), the belief that making personal connections to the material can be beneficial (nature of knowing and learning). This same belief is one of those two previously mentioned beliefs tested by question 12. Question 1 grouping with 12 but not loading well onto factor three is an indication that the items on factor three are indeed exploring some aspect of source of ability to learn, which is discussed shortly. Last, question 16 data indicates this same repeating belief that science can be learned, only this instance has students focusing on the need for hard work as a means by which to learn. Keep in mind that, in their written responses, these students favor hard work as being dedication and perseverance toward understanding the material, understanding likely developed through the more detailed insight given above with the written question response patterns for 5 and 11.

This factor best represents a student's belief regarding whether or not science can be learned and aligns up quite well with source of ability to learn. There is also commonality with "source of knowledge" present in epistemic literature, in that the knower is a constructor of meaning [20, 120, 135, 165]. Epistemologically, this factor seems to measure students beliefs for the efficacy of hard work and good study habits.

4. Factor 4: Nature of knowing and learning

The fourth factor in the five-factor theoretical model has only two items, 28 and 30. Question 28 appears to do an acceptable job of probing student views regarding the tentativeness of scientific knowledge and its limitations. From a larger perspective, however, many students' responses were focused around the view that second-hand data opens up room for interpretation. This question (28) asks students to discuss why multiple theories for the fate of the dinosaurs exist, a context which students primarily consider in their answering of the question. Numerous written responses reveal a student idea that the large time gap involved obscures data and/or that since humans were not around at the time, we simply cannot know for certain. On question 30, students largely agree with the argument that "[...] it is possible to get the right answer without really understanding what it means." Their justification favors both a sophisticated view where they may be missing out on conceptual understanding and a more unsophisticated view in that they are missing out on understanding the algorithm necessary to solve the problem. It is also of note that a non-negligible number of students agreed with Jessica, stating that it may be best to simply move on in the interest of time (as opposed to correctness, which the question intended to focus on).

The most common theme between these two items is that a number of conclusions may be achieved despite having uncertainty in the process driving that conclusion. Based on written response data, students believe there to be valuable knowledge in evaluating the process by which conclusions have been drawn. Questions 28 and 30 thus appear to be gauging the value a student places on evaluating current knowledge in the context of science, particularly when students are aware of possible ambiguity in the process by which that knowledge was acquired. In question 28 students evaluate the knowledge put forth by the scientific community whereas in question 30 students evaluate their own knowledge. With respect to the EBAPS axes, these questions appear to be centered on the nature of knowing and learning (see Table A.1). There are also similarities to a justification of knowledge component within the epistemic theories of both King and Kitchener and Kuhnwhere knowledge requires no justification and there is an acceptance of facts as compared to critical evaluation of the knowledge and judgments to which one is exposed [120, 125].

As an aside, it was seen that question 27 occasionally had a strong loading with this factor as well (but struggled with issues of low commonality). Beyond just strong loadings, there is a case to be made for item 27 being considered on this axis. A portion of the argument in item 27 stands out, specifically in the pseudostudent Julia who states, “I still think science applies to almost all real-world experiences. If we can’t figure out how, it’s because the stuff is very complicated, or because we don’t know enough science yet.” No written response data exist for question 27, but Julia’s statement represents one with which factor four is familiar: evaluation of knowledge. In analyzing items 28 and 30 students are reflecting on how the current state of knowledge came to be, assessing both personal (such as question 30) and external (such as question 28) conclusions. Question 27 upholds the theme of acknowledging uncertainties and their role in the current state of knowledge and knowing.

The relationship of items 27, 28, and 30 to the work of King and Kitchener in their reflective judgment model is intriguing enough to be expanded upon in slightly greater detail. The reflective judgement (RJ) model outlines key stages in the development of critical thinking skills in the context of "ill-structured" problems, or problems which lack an absolute truth and cannot be approached algorithmically [120]. Each stage within the model attempts to link epistemic views to decisions (judgments) made for ill-structured problems. Keen are King and Kitchener on "epistemic cognition," which pertains to how individuals think and reason in relation to their beliefs of "[...] the limits of knowing, the certainty of knowing, and the criteria for knowing" when encountering these types of problems [121]. In written responses to the data, students frequently displayed aspects of all three primary domains in the RJ model: pre-reflective thinking, quasi-reflective thinking, and reflective thinking [120]. The most frequently occurring responses align with the quasi-reflective domain where students often made statements of "missing knowledge" and "uncertainty" within the knowledge claims, leading them to have individualistic choices of evidence in supporting their conclusions. It is possible that the items pertaining to factor four are measuring the sophistication of a student's epistemic cognition. This is satisfactory, as metacognition alone could not account for how students were viewing and working with knowledge in their answering of items 28 and 30 for the written response data.

5. Factor 5: Quick learning

Questions 4 and 9 make up the fifth and final factor of the five-factor theoretical model. Recall how in testing stability of the factors that there would arise inconsistencies between group A and group B for items 4 and 9. Namely, group

A had 4 and 9 load alone on a factor together whereas group B saw item 9 load onto factor three while item 4 represented its own factor of one item. From simply reading question 4 it is possible that it is assessing quick learning (that science is learned quickly or not at all), like that of Schommer [183,184]. Although no interview data of item 4 are available, we do have data regarding student definitions of natural ability, a word which question 4 (or more specifically, quick learning) has some ties with. To students, one of the primary traits of natural ability is for a student to grasp material in a much shorter time than others. The term "natural ability" appears along with "learn" in question 9 and as such, it is plausible that some students may also view question 9 as stating "Someone who does not learn the material quickly can still learn the material well [...]." Hence, it would be possible that question 9 is probing quick learning as well as the belief to which factor three belongs (the degree to which students believe they can learn science by utilizing hard work and good study habits).

Recall that for group A, item 16 was seen to group with 4 (and 9) as well for the five-factor model. In responses to item 16 (which belongs more so to factor 3, source of ability to learn) most students were observed to state that given enough time, anyone can learn science. However, a noticeable portion of these responses revealed the belief that learning occurs at different rates for different people. This belief is not necessarily exemplifying a belief in quick learning and brings into question what items 4 and 9 may be testing. Epistemologically, factor five could be assessing the extent to which the ability to learn the material quickly affects whether material can be learned at all. However, this factor could also merely be assessing the extent to which they believe that learning may occur at different rates for different people, but inevitably can occur. There is currently an effort being made to acquire written response data on item 4 prior to the CFA analysis to clarify this ambiguity. Although this factor appears to be our weakest factor, and could be dropped within CFA, it is retained as it appears to be measuring a belief which is prominent and distinct within students.

Theoretical Three-Factor Model The three prominent factors belonging to the five-factor theoretical model (factor one, factor two, and factor three) are essentially the same factors as are present in the three-factor theoretical model, and thus assess those same beliefs. The only difference is that in the three-factor model question 9 is now grouped with factor 3, which probes *source of ability to learn.* Question 9 did assess *quick learning* (or a similar construct), but also assessed the ability to learn science, i.e., *source of ability to learn.* This could be seen in written response data where students seemed to focus their answer on the latter portion of question 9 and thus read it as "Someone [. . .] can still learn the material well even in a hard chemistry or physics class." With question 4 no longer present in the analysis (due to restriction to three factors and process of item removal), question 9 consequently no longer experiences a cross-loading associated with quick learning. Because of the involvement with another epistemic construct, it is debatable whether or not question 9 should even be considered for this factor in the three-factor theoretical model. This is something to be explored further in future work involving confirmatory factor analysis.

A Sixth Factor Mentioned previously was the presence of a possible six-factor model, this included another factor beyond that of the five-factor model and would be composed of items 1, 12, and 26. Written response data with question 1 revealed that students saw relating science to their own experiences as beneficial. A similar response pattern was seen for question 12, in which students responded with the belief that relating science to their own ideas was beneficial. Question 12 also had written responses which exemplified the items association with factor three, in that students believe there are a variety of ways in which one can learn the material. No written response data are available for question 26, however, reviewing the item can still allow for some conjecture regarding what it measures. In particular, the response of pseudostudent Justin in question 26 (that putting science concepts into his own words assisted his learning) highlights a common theme across all three items (1, 12, and 26): that relating aspects of science to ones self is beneficial to learning. It is possible that factor six is outlining a specific study or learning habit that may be implicit in factor three.

Taken together, it is believed the items associated with factor six measure the value students place on making personal connections with science. The stability of this factor is brought into question though, given that question 12 also belongs to factor three and that the other pseudostudent Dave (from question 26) is likely putting emphasis on a separate epistemic belief known as source of knowledge (what does the student depend on for knowledge, themselves or an authority figure?) [104] .

The EBAPS in Astronomy

In this section the axes of the EBAPS (Table A.1) will be discussed in regards to their relationship with the factors found during the exploratory factor analysis. We assert that this is by no means a rigorous test of the validity of the EBAPS instrument or the items as they appear on the axes of the EBAPS. This is an exploration of the prominent epistemological beliefs introductory astronomy students are exhibiting when responding to the EBAPS and how those beliefs may be related to what the EBAPS claims to measure.

The first EBAPS axis, the *structure of scientific knowledge*, mirrors very closely the structure of science factor revealed within this study. The authors of the EBAPS claim the items pertaining to axis one measure the extent to which knowledge in science is viewed as disjoint or isolated, as compared to a unified whole. In truth, the *structure of science* is in many ways measuring what structure of scientific knowledge claims to measure. The conscious choice to not perfectly align this factor with axis one of the EBAPS was due to student written response data. These data revealed that students were drawing their views of the cohesiveness of scientific knowledge from principles related to the nature of science. Evidence hints that this factor may be measuring student sophistication of NOS tenets by proxy with EBAPS questions from structure of scientific knowledge, "real-life applicability," and "evolving knowledge."

The nature of knowing and learning is axis two of the EBAPS and measures the value a student places on constructing their own knowledge, as opposed to passively receiving knowledge. This axis probes a variety of methods regarding how a student constructs their own knowledge, from metacognitive strategies to personal experience. It would seem the factor *nature of knowing and learning* partially tests these same traits. Mentioned before, *nature of knowing and learning* may be more aptly called *justification of knowledge*, as this factor is an indicator of how students reflect upon what is "known" and how it may, or may not, be known. The possible sixth factor put forth in this study, what we would call *personal connections*, may also fit with this axis. *Personal connections* appeared to be a factor whose questions explore the extent to which students value making personal connections to science. These personal connections to science involve experiences, ideas, and/or words pertaining to the self. Similarly, nature of knowing and learning claims to test, in part, students ability to relate "new material to prior experiences, intuitions, and knowledge."

Axis three of the EBAPS is real-life applicability. It is believed that no one factor aligns well with this axis, which claims to measure the extent to which science is applicable outside the classroom or laboratory; how well science applies to real life. *Structure of science* would have the most in common with this axis, as students actively discussed (via written response) science in everyday life for questions along that factor. In particular, question 14, where many student responses viewed science as the machinery behind everyday life (thus politicians should have a working knowledge of science). Again, *structure of science* was not testing real-life applicability but aspects of the nature of science, yet these NOS principles would often come forth in real-life issues concerning science.

Evolving knowledge, the fourth EBAPS axis, seeks to measure the extent to which students exhibit aspects of dualistic views or relativistic views with scientific knowledge [165]. Dualistic views hold all knowledge as being either right or wrong (absolute truths exist), whereas relativistic views hold scientific knowledge as purely subjective with no absolute truths. Not often do written responses from students take a stance of science as absolute truth which did not change, nor did they frequently exhibit complete immersion in subjectivism, unable to appreciate any evidence put forth by science. Despite not observing these extremes within the data, it is likely that they have some influence, as epistemic literature has essentially confirmed their presence [20,120,125,165]. The apparent lack of these epistemic perspectives emerging as a factor within this study may be due to the relatively few items which explicitly address them (6, 28, 29) and/or because they are extremes which need to be analyzed independently on a collection of separate questions.

The fifth and final axis of the EBAPS is source of ability to learn. This axis claims to probe students beliefs that they can become better at science, focusing on hard work over natural ability as well as the use of good study strategies. EFA revealed that this axis is seen as two distinct axes according to the students. Specifically, views about the "efficacy of hard work and good study strategies" seems to be getting tested by the factor *source of ability to learn* (whose name was, of course, motivated by its EBAPS counterpart). However, the first part of the EBAPS authors' description of the fifth axis "Is being good at science mostly a matter of fixed natural ability? Or, can most people become better at learning (and doing) science? [through hard work]" appears to be linked to its own factor, which was called *innate ability vs hard work*.

To summarize, structure of scientific knowledge, nature of knowing and learning, and source of ability to learn all appear to play a respectable role within the factors revealed in the exploratory analysis. Evolving knowledge did not appear to play a prominent role, but likely this is due to question limitation or the need for independent testing of its constructs (absolutism and relativism). Real-life applicability, as defined by the authors (Table A.1), was the only axis that did not hold much significance in student response patterns to the EBAPS. This could be because the trait was playing a more subtle role than could be seen in EFA and, in some respects, the written response data, or because it is better accounted for with other fundamental epistemic constructs.

Conclusions

When exploring the underlying structure of student responses to the EBAPS survey in an introductory astronomy class, there were seen to be three to five notable factors explaining the relationship between item responses. A three-factor and five-factor theoretical model were discussed and their factors placed into context utilizing data from student responses to EBAPS items. Both theoretical models shared three factors, what we called *structure of science*, *innate vs hardwork*, and *source of ability to learn*. *Structure of Science* appears to represent the extent to which students have sophisticated views regarding the nature of science, in particular the philosophical and psychological tenants of NOS [142]. *Innate vs hard work* directly confronts the views of students regarding the influence that natural ability has as compared to hard work on an individual's ability to succeed. *Source of ability* to learn is best described by the EBAPS itself in that the items belonging to this factor "probe students epistemological views about the efficacy of hard work and good study strategies, as distinct from their self-confidence and other beliefs about themselves." [68]. The two additional factors within the five-factor theoretical model are *nature of knowing and learning* and *quick learning*. *Nature of knowing and learning* is exploring the value that a student places on justifying what they know and the acknowledgment that learning can occur through this justification. This factor also shares a notable relationship with what King and Kitchener call "epistemic cognition" [121]. Last, there is *quick learning*, which probes the extent to which students believe they can still learn the material despite not being able to grasp it as quickly as their peers with natural ability. We do note that it is possible that this final factor is exploring a belief that learning occurs at different rates for different people, but can occur. Future works hopes to obtain a clarification regarding this factor and what it is measuring.

In analyzing student responses to EBAPS items, themes regarding more fundamental epistemological beliefs were seen involving the nature of knowledge and the nature of knowing. Beliefs such as certainty of knowledge, the simplicity of knowledge, and the justification of knowing [104]. To speculate, response data hint that these core epistemic beliefs influence, if not generate, several of the factors in this paper which pertain to the nature of learning and the nature of intelligence.

Future Work

Future work will involve the use of confirmatory factor analysis in testing the hypothesized variable relationships and the fit of the proposed models. The factors as put forth by the authors of the EBAPS will also be subjected to CFA. A more detailed question-by-question analysis of the EBAPS will be performed. The question-by-question analysis is intended to elaborate more extensively on how students in the astronomy course viewed and responded to questions on the EBAPS and how further clarity and orthogonality could be obtained on these items within this context (introductory astronomy).

End of previously published material as adapted from [116].

Development of an Epistemological Response Framework to the EBAPS

Introduction

An exploratory factor analysis (EFA) is also done on the EBAPS utilizing the R programming software to refine the results found within previous SPSS EFA work (See Ch. 4 on page 83) [116]. The analysis will assist in committing to a model (or set of models) which will then be tested within confirmatory factor analysis. The results of the previous SPSS EFA work will be frequently considered during the analysis. Furthermore, a unique approach will be taken within "R", for which emphasis will be placed on the partial correlations between items so as to help uncover the most prominent epistemic factors influencing student responses to the EBAPS. These findings will be integrated with student written responses (Appendix F) when applicable in order to identify a plausible epistemological framework for student responses to all 30 questions on the EBAPS.

Data utilized within were essentially identical to those present in the SPSS exploratory factor analysis [116]. That is, the data represent all the students from section one of the introductory astronomy course which took place from fall of 2012 through spring of 2017. Furthermore, only the post-test data from the EBAPS are used. The entirety of the analysis is done with the "R" programming language under the "RStudio" integrated development environment [4,5]. Within Appendix B, descriptive statistics for the data may be found in Table B.1 and measures of sampling adequacy (MSA) values may be found in Table B.2.

Working Towards an EFA Model

As seen in previous EFA work (Ch. 4 on page 83), it was found that there are three prominent factors within the data, with up to six factors in all. It was also

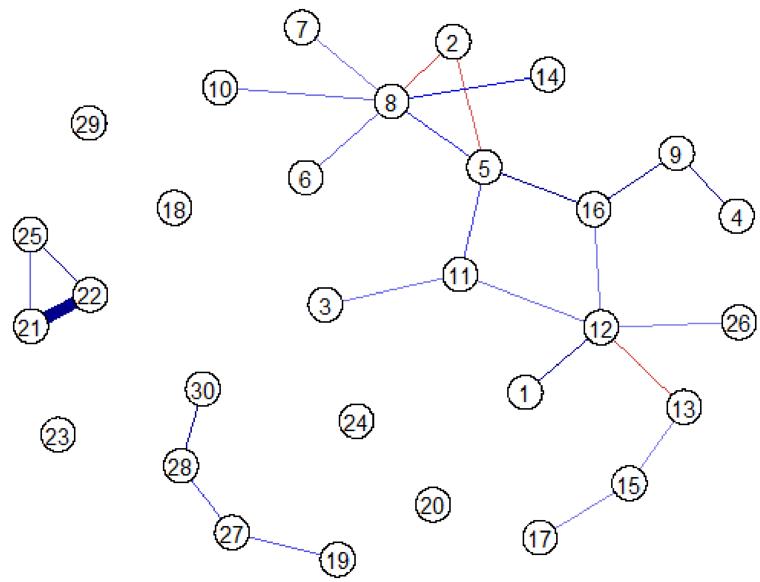

Figure 4.2: Map of partial correlations between items on the EBAPS for all Section 1 data. Strength of correlation indicated by width of line, direction of correlation indicated by color (blue-positive, red-negative). Items connected by a line indicate significance at the $p < 0.100$ level.

seen that each subsequent model appears to build off the previous (additional factors appear while previous remain). The refinement of factors within this work depends heavily on the partial correlations present between items within the data. With this in mind, the refinement of a factor model will begin with a visual representation of the strength, significance, and direction of the partial correlations between items within the analysis.

Figure 4.2 is a visual representation of the partial correlations between all items within the data. A line connecting items indicates that the partial correlation between those items is significant at the $p < 0.10$ level (Bonferroni corrections applied). This level was chosen as it was observed to be a good starting value for filtering items, given that it well replicated item importance within initial filtering results of the

SPSS EFA analysis (this p-value implied that items 18, 20, 23, 24, and 29 should be removed, criteria in agreement with both the researcher and previous results as seen in Ch. 4 on page 83). Furthermore, it is not desired to create too strict of an initial filtering criteria as item significance will eventually be further refined through the use of student written response data. Continuing, the thickness of the line is an indicator of the strength of partial correlation between the items. Lastly, the color of the line indicates the direction of the correlation, with blue representing a positive relationship and red representing a negative relationship.

The general philosophy behind refining the data and creating factors was as follows:

1. Assess the partial correlation map to assist in determining which initial items to remove from EFA. Items without a line (i.e. items that do not share even one significant partial correlation with any other item) are removed.

2. Conduct an EFA to determine a single primary factor present within the data.

3. Assess the partial correlation map using just the items whose loadings were greater than 0.39 on the primary factor. The ideal configuration is one in which every node (item) in the plot is connected via a line to every other node (i.e. that all the items present within a factor share a significant partial correlation with every other item in the factor). Remove any items which do not contribute to an "ideal configuration". The first item to be removed is the one with the fewest connections (lines) to other items. If there is a tie, repeat the factor analysis with just the items from this factor and remove which ever of the tied items has the lowest communality. Repeat this process until an "ideal configuration" is achieved.

4. Return all items to the data set except those removed in step 1. Conduct another EFA to determine the two most prominent factors present within the data.

5. Assess the partial correlation map using just the items whose loadings were greater than 0.39, cross-loadings are allowed so long as the difference for the items primary and secondary factor loading values is greater than 0.2 [109]. Furthermore, under the aforementioned loading criteria, no new items are to belong to an established factor (such as the factor put forth in step 3) and as such any items which do so are not to be brought forward onto the map.
   In the case of multiple factors an ideal configuration will not have two or more nodes which belong to a single factor as also having connections to a common secondary factor. If this is the case, remove the nodes (EBAPS questions) which display this behavior individually. Whichever resulting map no longer creates the aforementioned violation is to be preferred (if the violation still exists, there are multiple ways in which it could be handled. As it did not occur within the analysis, it is not discussed here). Once this is resolved, attempt to recreate an ideal configuration within the new factor as defined in step 3.

6. This basic process is repeated until factor analysis no longer yields factors with three or more items, or if it is no longer possible to create a scenario in which an "ideal configuration" is attainable.

During factor analysis the same factor extraction method (principal axis factoring) and factor rotation method (direct oblimin) utilized in the SPSS analysis is employed within "R". The refinement philosophy outlined above will assist in determining an optimal number of factors to retain and also help determine the fundamental items belonging to each factor. As previously explored in SPSS (4 on page 83) , between three and six factors should be expected.

Constructing the Factor Model

Following the philosophy outlined just above, a one-factor model was constructed. The initial map of partial correlations (Figure 4.2) resulted in the permanent removal of items 18, 20, 23, 24, and 29 from the analysis. Upon running a one-factor EFA on the remaining items, it was seen that items 4, 5, 6, 8, 9 ,10, 12, 14, and 16 were all present on a single factor, with loadings greater than 0.39 [91]. Another partial correlation map was created which contained only the aforementioned items, revealing items 4 and 6 to have the fewest connections, with three each. An EFA was run on the remaining items (4, 5, 6, 8, 9 ,10, 12, 14, and 16) which revealed item 4 to have a lower communality than item 6 and was thus removed from the analysis. This basic process was repeated, resulting in the subsequent removal of items 10, 6, and 14, respectively. Figure 4.3 represents the remaining partial correlation plot.

It is now said that the one-factor model is displaying an “ideal configuration” in that all items share significant partial correlations with all other items.

Next, a two-factor model was sought. All items present in the original analysis (i.e. all items except 18, 20, 23, 24, and 29) were reintroduced into the data and a two-factor EFA was conducted. Again, items 5, 6, 8, 9 ,10, 12, 14, and 16 were all present on a single factor. As this factor pertains to the original primary factor, items 6, 10, and 14 were not brought forward in the mapping analysis (these items have already demonstrated that they do not contribute to an “ideal configuration” for factor one, the original factor). For factor two, items 21, 22, and 25 all had loadings greater than 0.39 and were thus brought forward. The resulting partial correlation map containing the items brought forward needed no additional refinement, as an

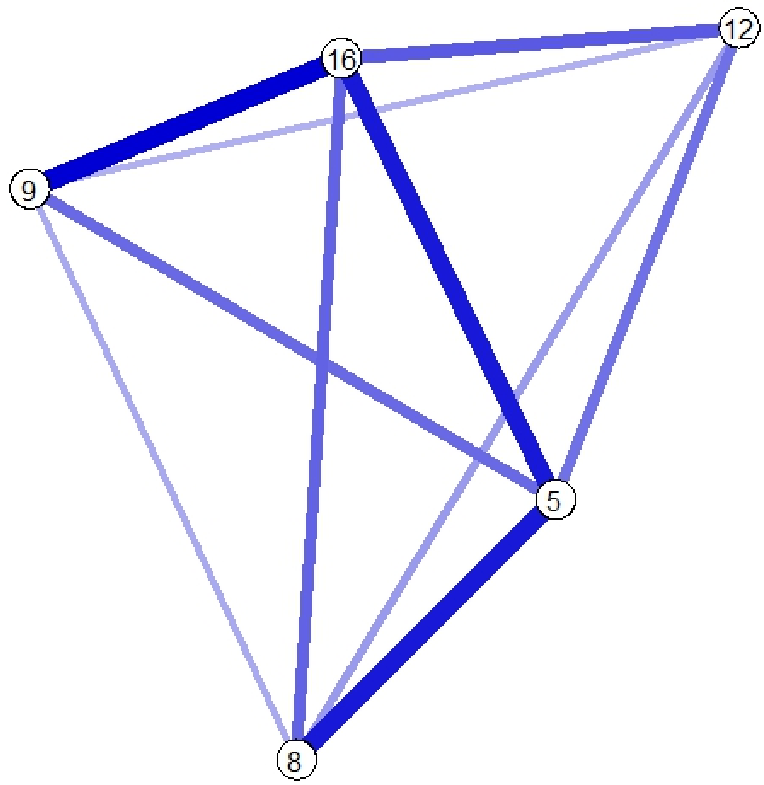

Figure 4.3: Map of partial correlations between items on the EBAPS for all section 1 data, after filtering was applied. This mapping represents the items on a single-factor EFA. Strength of correlation indicated by width of line, direction of correlation indicated by color (blue-positive, red-negative). Items connected by a line indicate significance at the $p < 0.100$ level.

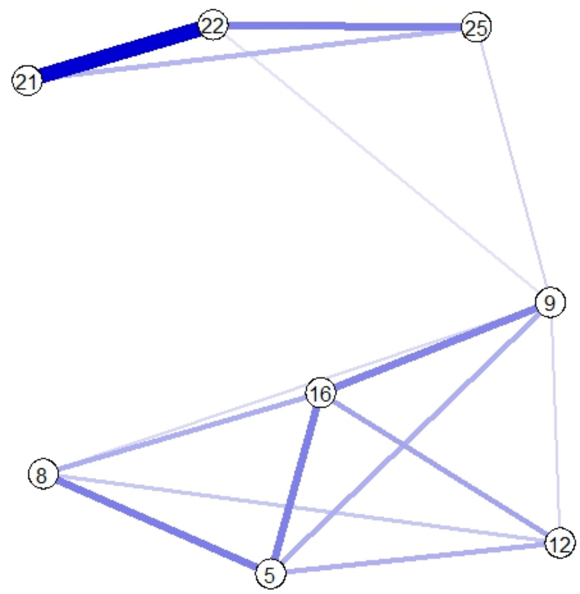

Figure 4.4: Map of partial correlations between items on the EBAPS for all Section 1 data, after filtering was applied. This mapping represents the items on a two-factor EFA. Strength of correlation indicated by width of line, direction of correlation indicated by color (blue-positive, red-negative). Items connected by a line indicate significance at the $p < 0.100$ level.

"ideal configuration" happened to already be present (see Figure 4.4).

A three-factor EFA was performed on all items present in the original analysis. This again resulted in a factor with items 21, 22, and 25, which had no other loadings greater than 0.39. The original factor of items 5, 9, 12, and 16 was present, however, item 8 moved on to the new factor. Item 11 had a loading greater than 0.39 on this original factor, but was not brought forward for mapping analysis (general rationale presented earlier). The newest factor contained items 6, 7, 8, 10, 14, and 15 with loadings greater than 0.39. The map containing the items brought forth for this set of three factors can be seen in Figure 4.5.

Note that this mapping does not demonstrate an "ideal configuration" in several

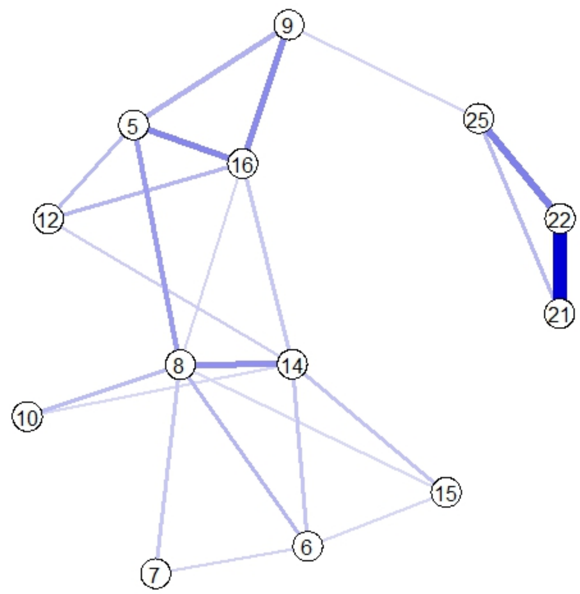

Figure 4.5: Map of partial correlations between items on the EBAPS for all Section 1 data, after partial filtering was applied. This mapping represents the initial items present on a three-factor EFA. Strength of correlation indicated by width of line, direction of correlation indicated by color (blue-positive, red-negative). Items connected by a line indicate significance at the $p < 0.100$ level.

ways, firstly, in how items 8 and 14 share connections with items from the original factor (composed of items 5, 9, 12, and 16). As such, either item 8 or item 14 should be considered for removal. However, item 8 once belonged to the original factor and was not considered for removal. That is to say, any new factors that arise should not require a removal of items from the analysis that have already been established as being part of an “ideal configuration” for a factor. Item 14 was removed from the analysis, after which it was seen that no further cases of multiple item connections to a common secondary factor were observed (item 8 still held several connections to the original factor). Within the mapping items 10 and 15 were seen to hold only three connections (out of four possible, ideally). An EFA was conducted on the remaining items (5, 9, 12, 16, 21, 22, 25, 6, 7, 8, 10, and 15). The resulting communality for item 15 was lower than that of item 10, and was thus removed from the analysis. Figure 4.6 represents a new mapping with the remaining items, this mapping demonstrates an “ideal configuration” with no further modifications necessary.

The presence of item 8 connecting to items within the original factor is not a concern, as the difference in loading between item 8 on the new factor and item 8 on the original factor was greater than 0.20, while also simultaneously loading onto the alternative factor with a value less than 0.30 [109]. In principal, this may imply that item 8 be represented as a cross-loading within CFA, as it could be probing two different aspects of student epistemologies.

A four-factor EFA was then performed on all items present within the original analysis. The four-factor EFA revealed a new factor containing items 27, 28, and 30, all with loadings greater than 0.39. As per usual, items which were seen to load onto already established factors were also brought forward for mapping while others which have previously been observed not to contribute to an “ideal configuration” for their respective loading factor were not brought forward.

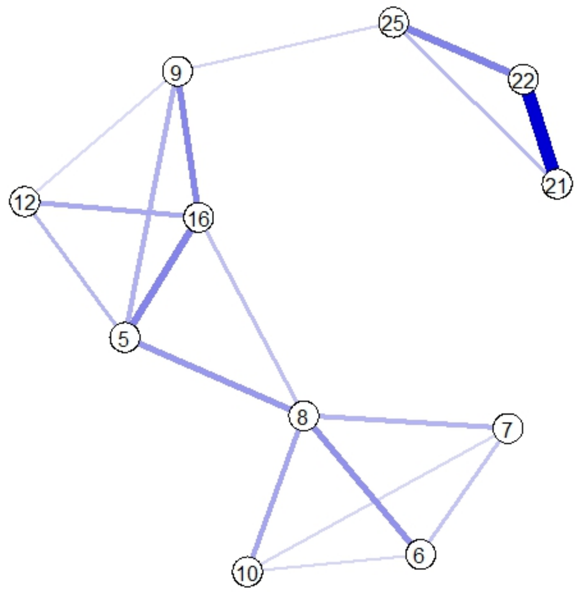

Figure 4.6: Map of partial correlations between items on the EBAPS for all Section 1 data, after filtering was applied. This mapping represents the final items present on a three-factor EFA. Strength of correlation indicated by width of line, direction of correlation indicated by color (blue-positive, red-negative). Items connected by a line indicate significance at the $p < 0.100$ level.

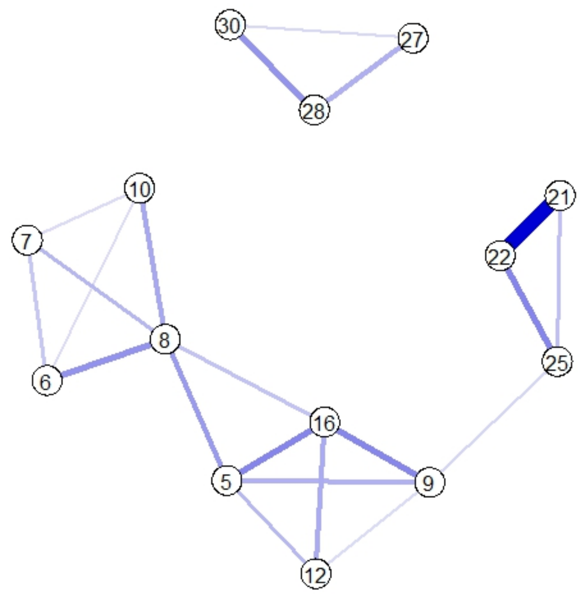

Figure 4.7: Map of partial correlations between items on the EBAPS for all Section 1 data, after filtering was applied. This mapping represents the final items present on a four-factor EFA. Strength of correlation indicated by width of line, direction of correlation indicated by color (blue-positive, red-negative). Items connected by a line indicate significance at the $p < 0.100$ level.

Table 4.9: The Proposed Four-Factor Model for Section 1 EBAPS Data. This model was achieved primarily through the use of a unique approach as outlined in Ch. 4 on page 115

Factor 1	Factor 2	Factor 3	Factor 4
Question 5	Question 21	Question 6	Question 27
Question 9	Question 22	Question 7	Question 28
Question 12	Question 25	Question 8	Question 30
Question 16		Question 10	

Figure 4.7 represents the mapping for the items brought forward, as this network of items demonstrates an “ideal configuration” no further modifications are necessary.

A five-factor EFA model was performed on all items present within the original analysis. This EFA did not reveal a reliable fifth factor, as the items belonging to the fifth factor (items 1, 12, 13, and 26), only item 12 had a loading greater than 0.39. It is then determined that a five-factor model is likely not to be a stable model due to this factor.

It follows that the results obtained while employing an exploratory factor analysis combined with filtering criteria reliant on a partial correlation mapping yielded a reliable four-factor model. This reliability takes the form of what has been referred to as achieving an “ideal configuration” for all factors within the analysis.

A crucial part of this analysis relied heavily on the researcher’s definition of “ideal configuration”. Precisely why the researcher desires this type of configuration and why particular decisions are made are outlined in more detail hereafter. First, the desire for all items within a factor to share a significant partial correlation is discussed. Should an item, such as item 29 in Figure 4.2 on page 116, not have a significant partial correlation with any other item then it may be viewed as unique. A unique item is independent, no other items are measuring what that item is measuring (this assumes that the item is actually measuring something and isn't, say, just a poorly worded item). If two items share a significant link (partial correlation), then at least one aspect from each item is measuring something unique, even after all other variables are considered. It would follow that if three or more items all share significant links among each other, that the collective set of items are making a unique measurement. Within an EFA framework, it could then be said that there exists a construct which is responsible for all these items being influenced similarly.

Second, the decision to remove the weakest item (the item with fewest number of connections) one at a time follows similar logic. Any item which does not share connections with all other items (within a proposed factor) violates the tenet that these items are all being influenced by an underlying construct. Hence, the item displaying the fewest connections is considered the weakest link in the network and is removed first. In the instance that two or more items are the "weakest links", then EFA was utilized in order to view the communalities of the items. The communalities are a valued numeric quantity because they are related to error measurement (the larger the communality the lower the measurement error for that item). During EFA computation (Ch. 3 on page 56), the uniquenesses are effectively the differences between an item's actual correlation with itself and the item's recreated correlation with itself, a recreation dependent on the other items involved in the EFA. The communality for an item is one minus the uniqueness for that item. Thus, the larger the communality, the better that item is at measuring what the other items are also attempting to measure (less error). Finally, during the refinement process should the removal of an item cause the network as a whole to become less "ideal", and additional item removals do not remedy this, then it may be best to terminate the factor attempting to be found and deem the factor as unreliable. This was nearly a concern during the attempted refinement of the third factor, but otherwise not encountered during this process.

The last key decision involves a scenario in which two (or more) items from a new factor are seen to link to two or more items from an established factor. This occurred for items 8 and 14 during the discovery/refinement of factor three. It was proposed that items exhibiting this behavior be removed individually, and whichever items removal creates the more "ideal" factor indeed be the item that gets removed. This decision is based on the researcher preference that factors be as independent as possible (to assist in making teh description of factors as simple as possible). In the case of items 8 and 14, item 8 had already been seen to belong to factor one. Item 8 was not considered for removal because additional factors should maintain independence from already constructed factors. Thus, the need to remove item 8 in order to achieve independence compromises this philosophy. Should this philosophy be ignored while maintaining the "weakest link" philosophy, the resulting three-factor partial correlation network is shown in Fig. 4.8.

Theoretical Response Structure

The following will include a discussion of the EFA results yielded by SPSS work (Ch. 4 on page 83) as they compare to the above results given within R. Note that while the fundamental EFA calculations were the same in both SPSS and R (principal axes, oblique rotation, etc.), the manner in which factors were uncovered was mostly unique between the two. Rationale for the linkings present within the partial correlation mapping seen in Fig. 4.2 will also be expanded upon. All written response data referred to will pertain to that which is presented within Appendix F.

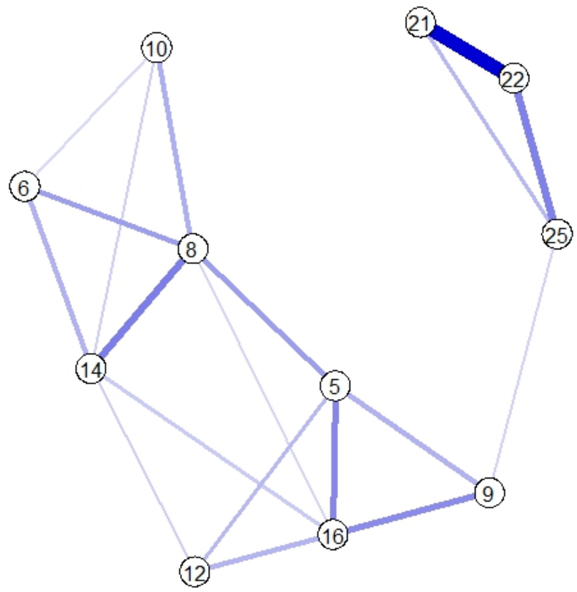

Figure 4.8: Map of partial correlations between items on the EBAPS for all Section 1 data, after an alternative filtering was applied. This mapping represents the final items present on a three-factor EFA. Strength of correlation indicated by width of line, direction of correlation indicated by color (blue-positive, red-negative). Items connected by a line indicate significance at the $p < 0.100$ level.

Table 4.10: Summary of factors found within an EFA analysis utilizing unique approaches as described in Ch. 4 on page 83 and Ch. 4 on page 115

Factor 1		Factor 2		Factor 3		Factor 4		Factor 5	
SPSS	R	SPSS	R	SPSS	R	SPSS	R	SPSS	R
5	5	21	21	6	6	—	27	—	1
9	9	22	22	—	7	28	28	4	—
11	—	25	25	8	8	30	30	9	—
12	12	—	—	10	10	—	—	—	12
16	16	—	—	14	—	—	—	—	26

Comparing SPSS EFA Results and R EFA Results Factors and their corresponding items as found within the SPSS methodology and the R methodology are presented side-by-side in Table 4.10. Note that in SPSS a three-factor model was also proposed, which would be represented by the factors 1, 2, and 3 in Table 4.10. Furthermore, within R, a four-factor model was settled on and is represented by factors 1, 2, 3, and 4. Factor 4 and 5 both represent factors as found in the SPSS five-factor model, while in R factor five represents an unstable factor that was not considered for model inclusion (but is presented here for discussion and comparison purposes). The purpose of this comparison is to finally settle on a primary model for use in confirmatory factor analysis (CFA). Overall, the results would indicate that a four-factor model may be the best choice to commit to, given that these two different approaches yielded such similar findings up until the fifth factor.

Factor 1 from Table 4.10 contains items 5, 9, 12, and 16 in both SPSS and R. SPSS also includes item 11, so it must now be determined if item 11 should be included in the final version of this factor. Within SPSS this factor was described as "[...]

a student's belief regarding whether or not science can be learned [...]" [116]. The common trait connecting items 5, 9, 11, 12, and 16 are views about study/learning strategies and their effectiveness. However, there appears to be a subtle difference between item 11 and items 5, 9, 12, and 16. Student written responses to item 11 focus on study/learning strategies as they pertain to identifying whether or not they understand the material. There is an active component to item 11 responses which is heavily associated with metacognition. Even the most prominent coding from the neutral/unsophisticated responses to item 11 indicates this, with students referring to the belief that they need to talk with other students in order to know how they did on an exam. In items 5, 9, 12 and 16, students are expressing how well they believe they can do (as dictated by hard work and study/learning strategies), while item 11 has students reflecting on how well they are doing (as dictated by metacognitive ability). Item 11 is believed to have grouped with items 5, 9, 12, and 16 because in students reflecting upon how they are performing (metacognition) they are often considering how they could have performed differently (as dictated by hard work and study/learning strategies). As item 11 appears be measuring what items 5, 9, 12, and 16 measure by proxy of metacognition, it is decided that item 11 not be included with Factor 1. The definition of this factor as originally posed in the SPSS work remains unchanged despite the removal of item 11. That is, items on this factor would seem to be measuring the efficacy of hard work and study strategies [116]. This factor (containing items 5, 9, 12, and 16) is to be called *source of ability to learn* because it aligns up exceptionally well with the EBAPS axis of the same name. One curiosity from analyzing item 11 is whether or not metacognitive ability is largely a reflection of a student's sophistication along this factor. To clarify, a student's sophistication along *source of ability to learn* may hold a significant positive causal link to their metacognitive ability.

There is little to discuss about Factor 2 as results from both the SPSS and R analyses are identical (contain items 21, 22, and 25). As such, the definition as originally presented in the SPSS work will remain unchanged. Items on this factor "[. . .] pit the influence of hard work against that of natural ability when assessing the role that they play in student views of success" [116]. Sophisticated responses would have that hard work is most important in enabling success while unsophisticated responses would have that natural ability is most important (by comparison) in enabling success. Of course, there is still some ambiguity in what it means to be "successful" in either science or life. All things considered, this factor will be called *hard work vs. natural ability*.

Factor 3 within the SPSS work was originally believed to be probing "[. . .] student views with respect to the philosophical and psychological aspects of the nature of science" [116]. The presence of item 7 as having replaced item 14 within the R exploratory work demands that the nature of this factor be reconsidered. A common trend within student written responses to items 6, 7, 8, and 10 is the evaluation of potential gaps in knowledge. In items 6, 8, and 10, these gaps are typically reflected upon in student written responses with their views regarding the nature of science (NOS), which has been previously expanded upon in Ch. 4 on page 83. Within item 7, the students often rely on placing themselves in the position of the instructor. In doing so they openly acknowledge that teaching can improve learning, as the students view it as a type of study/learning technique. Specifically, the technique of actively working through the material and being subjected to student questioning. Thus it may be said that both item 7 and items 6, 8, and 10 all assess some aspect of identifying/evaluating gaps in knowledge. However, additional insight is provided by a code which was prevalent among the unsophisticated written student responses to item 7. This code was associated with the student belief that teaching requires mastery prior to class. This is compared to the written responses of item 10, in which there was a strong overtone of trust-in-science mindset. It is with these considerations, combined with the previous SPSS work, that it is proposed items 6, 7, 8, and 10 are probing student views about authority. Specifically, their views about authority as either all-knowing or as active learners who do not hold all the knowledge about the subject. Factor 3 is believed to be best represented by items 6, 8, 10, and 14 (not 6, 7, 8, and 10). If 6, 7, 8, and 10 are truly testing views about authority, then students are viewing science itself as a type of authority when responding to items 6, 8, and 10. This is believed to be the case, however, science is being viewed as authoritative predominantly due to student views about NOS (as gleaned from written response data). Fundamentally, it is student views about science which appear to be influencing the written response patterns seen within items 6, 8, and 10. Sophisticated views about the tenets of science then elevate student views regarding the credibility and reliability of science, thus compelling students to treat science (e.g. scientists and/or scientific findings) as more of an authoritative figure. How students then treat this figure corresponds to their views as described above (authority as either all-knowing or authority as an active learner which does not hold all the knowledge). It is because these authoritative views (for items 6, 8, and 10) are believed to depend upon the student's sophistication in NOS that Factor 3 from the SPSS analysis is favored over Factor 3 from the R analysis. So this factor shall be called *structure of science*. Note that, in the most general sense, this factor could also simultaneously represent the extent to which students view science/scientists/scientific knowledge as being an authoritative figure.

Factor 4 from the SPSS and R analyses are nearly identical, with the exception of item 27 grouping with items 28 and 30 in the R analysis. The possibility of item 27 belonging with items 28 and 30 was already acknowledged within the SPSS analysis

(See Ch. 4 on page 83). At the time of the SPSS analysis, written response data regarding item 27 was not available. Now that written response data is available on item 27, it appears as though the item has three prominent themes. The first theme is the applicability of science outside the classroom. The second pertains to a view of science as difficult to apply, due to the number of complications involved (e.g. variables). Last, is the view of ambiguity in applying science as resulting from a variety of other uncertainties (simply put, not knowing enough information/science yet). It is believed that the last of these three themes is responsible for item 27 grouping with items 28 and 30 within the R analysis. As there now exists written response data for item 27 which links it to items 28 and 30, Factor 4 from the R analysis is taken to be favored over that of Factor 4 from the SPSS analysis. The definition of Factor 4 is kept the same as that which was stated in the SPSS analysis, in that these items seem "[...] to be gauging the value a student places on evaluating current knowledge in the context of science, particularly when students are aware of possible ambiguity in the process by which that knowledge was acquired" [116]. It is advised that the reader refer the corresponding section for this factor within Ch. 4 on page 83, which elaborates on how this factor relates to justification of knowledge and "epistemic cognition" [120, 125]. This factor will be called *evaluating uncertainty*.

Factor 5 from the SPSS work and R work were seen to be different. In SPSS, it was observed that this factor is composed of EBAPS items 4 and 9, for which no written response data existed for item 4 at the time. In the spring of 2018 students were able to provide written responses to question 4, allowing for more stable conclusions regarding the nature of what items 4 and 9 are testing. In the original SPSS work it was thought that this factor may be measuring "quick learning" (students believe they either learn the material quickly or not at all), a construct proposed by Schommer [183, 184]. The description of this factor as pertaining to "quick learning" is no longer believed to be the case. Instead, response data from items 4 and 9 indicate that the link connecting them pertains to a different aspect of natural ability. Question 9 response data has one clear focus, hard work. An aspect of this theme arises that focuses on the belief that lack of natural ability means a need for more hard work (i.e. effort/practice/studying/time). A primary response type in question 4 is the belief that some people learn faster than others, for a variety of reasons (learning style, past experience, motivation, etc.). Hence, it is proposed that factor 5 within the SPSS work (items 4 and 9) be measuring the extent to which students believe that learning occurs at different rates for different people. This could be called *learning rate*, for brevity.

Factor 5 from the R work contains items 1, 12, and 26. These items are believed to be linked by student views about whether or not relating science to self is beneficial. In item 1 this takes the form of relating personal experiences to science as a possible study/learning strategy. In item 12 students focus on the value of relating their own ideas to the text as being valuable to learning. Lastly, in item 26, students put forth the view that translating information (putting concepts into your own words) is valuable to learning. To summarize, the factor composing items 1, 12, and 26 shall be called *science to self*, as these items would seem to gauge the extent to which students believe making active use of connecting their own experiences, ideas, and beliefs to science can assist in their learning. Note that these items share a striking similarity with the EBAPS axis "nature of knowing and learning", and could very well be the same. If a fifth factor were to be committed to, *science to self* would be favored over *learning rate* due to the inconsistent views that students have regarding their definitions of "natural ability" (see Appendix F). To elaborate further, a student's sophistication within *learning rate* cannot be conclusively attributed to a belief in innate natural ability or some other construct as previously put forth (past

Table 4.11: The Final Proposed Four-Factor Model for Section 1 EBAPS Data. This model was achieved primarily through the use of multiple perspectives as outlined throughout Ch. 4

Source of Ability to Learn	Hard Work vs. Natural Ability	Structure of Science	Evaluating Uncertainty
Question 5	Question 21	Question 6	Question 27
Question 9	Question 22	Question 8	Question 28
Question 12	Question 25	Question 10	Question 30
Question 16		Question 14	

experience, motivation, etc.). As such, *learning rate* is likely a proxy measurement of a set of widely varying student views regarding natural ability. Further detail regarding a fifth factor from both the SPSS and R work can be found just below, where partial correlation connections are expanded upon.

Table 4.11 represents the final factor model intended for prominent use within CFA testing. An additional factor (*science to self*) could be included, which would contain the items 1, 12, and 26. The nature of this fifth factor may put too much stress on the model, as it is likely measuring a subset of what factor one is measuring (the efficacy of relating science to self as a learning/study strategy) and would pull item 12 from *source of ability to learn.* Table 4.12 shows the correlations between the four primary factors, as well as the factor *science to self.*

<u>Partial Correlation Mapping</u> To assist in developing a more complete framework regarding how students are responding to the EBAPS Fig. 4.9 will be employed, which shows the original partial correlation network for all 30 items on the EBAPS (for Section 1 students). Colored nodes are associated with the four primary factors settled on for CFA work. The following section links student written response data to the significant partial correlations that are observed. Partial correlations that connect the

Table 4.12: Factor Correlation Matrix Results of Final Proposed Four-Factor Model for Section 1 EBAPS Data, Including a Possible Fifth Factor. Source of Ability to Learn is Factor 1. Hard Work vs. Natural Ability is Factor 2. Structure of Science is Factor 3. Evaluating Uncertainty is factor 4. Factor 5 is Science to Self

Factors	1	2	3	4	5
1	1.000	—	—	—	—
2	0.30	1.000	—	—	—
3	0.65	0.20	1.000	—	—
4	0.36	0.35	0.39	1.000	—
5	0.25	0.02	0.33	0.19	1.000

nodes of the primary four factors will not be investigated as their connections should be strongly related to their factor descriptions, which have already been discussed above.

Independent Items The first aspect of this partial correlation mapping will be the lack of any connections to items 18, 20, 23, 24, and 29. Of these items, only 24 and 29 have student written response data. By inspection of items 18, 20, and 23, it would seem that the common trend between them is the heavy context placed on gauging understanding within physics (and chemistry). These questions would be difficult to answer if one has not taken multiple physics courses or is not currently within a physics course.

Item 18 in particular almost certainly requires that a student comprehend/appreciate the concepts involved in the question (Newton's second law, predominantly) in order to value whether or not it is "...a good question to think about?" [68]. As such, item 18 is likely independent because it is effectively asking students whether or not they have taken a physics class and done well in that class (there may exist

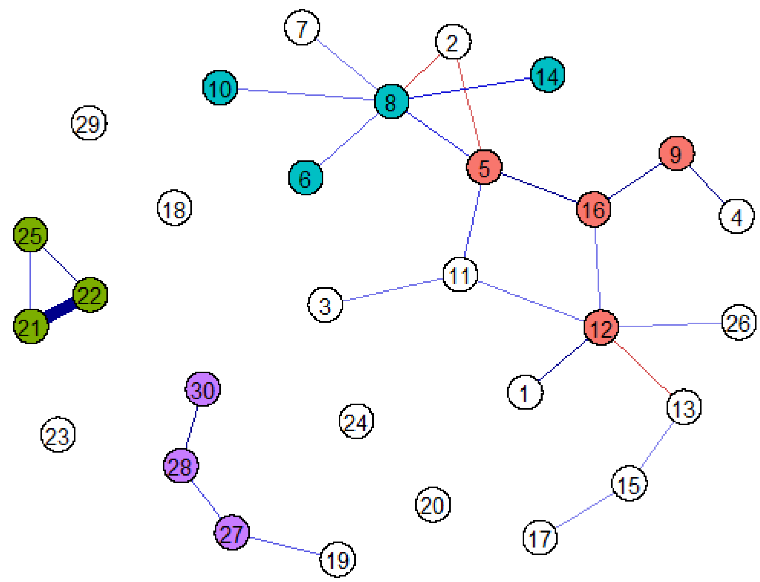

Figure 4.9: Map of partial correlations between all items on the EBAPS for all Section 1 data. Nodes which share the same color also share the same factor. Source of Ability to Learn is red. Hard Work vs. Natural Ability is green. Structure of Science is blue. Evaluating Uncertainty is purple. Strength of correlation indicated by width of line, direction of correlation indicated by color (blue-positive, red-negative). Items connected by a line indicate significance at the $p < 0.100$ level.

response variation within the problem which would likely represent random responses from individuals with little or no physics knowledge).

Item 20 attempts to probe how students relate formulas to concepts. Formulas themselves are mentioned twice within the EBAPS, question 17 and question 20. Within question 17 (discussed in more detail further below) students are addressing formulas in the context of solving a problem, whereas in question 20 the context places importance on relating concepts to the formulas. As such, question 20 may be independent of the other items because it is the only problem to explicitly do so.

Item 23 is similar to item 18 in that students who have little familiarity in physics are not likely to answer this question sophisticatedly. This question (should a test use a small number of long questions/problems covering several facts/concepts or a large number of questions covering individual facts/concepts) intends to put emphasis on the value students may place on connecting scientific knowledge. However, the feasibility of the aforementioned test philosophy plays a role in classrooms. Within the astronomy class of this study, the "unsophisticated" option is employed due to the labor involved in grading over 200 students. This alone could influence how students respond to the question, and thus gauge expectations as opposed to epistemologies. Item 23 is likely independent due to the lack of clarity between students expressing epistemic beliefs and those expressing expectations, as well as those who are simply uncertain because of the context (physics).

Until response data was acquired on item 24, it was difficult to understand why this item did not link to any other items. Upon parsing student written response data to item 24, there arose two overall problems with the question. The first (and less concerning) issue is that expectations were seen to occasionally play a role in student responses. To elaborate, some "unsophisticated" responses were justified by students remarking that science texts tend to separate chapters (and thus that is how it should be done). The notably more prevalent issue with item 24 is that it appears to be trapped between measuring "structure of scientific knowledge" and "nature of knowing and learning", as defined by the EBAPS [68]. Some students will claim that chapters should be linked because it makes the material easier to understand, an epistemic belief focused on learning. Other students will claim that chapters should be linked because science is interconnected, an epistemic belief focused on the structure of knowledge. Both types of students are responding in a sophisticated manner, but with a basis in different epistemological domains. It is for this reason that item 24 is likely independent, the underlying constructs governing the responses are varied and may thus be more aptly addressed by other problems. It may be more plausible, however, that this item is uniquely measuring the value that students place on linking knowledge within science.

Student written response data to item 29 also clarify why this particular item not be linked with any other items. The most predominant theme found in item 29 involves student views about theories and/or science. The most common views involved ideas such as "new information is capable of creating change", "questioning/testing what is known as a tenet of science", and "theories as being difficult to change". Those who answer the question in a sophisticated manner have similar arguments as seen to sophisticated responses within the proposed *structure of science* factor. However, responses were roughly equally distributed across all five response options. It is believed that this question is independent because it is the only question which directly addresses the tentativeness of scientific findings. Item 6, as posed by the authors of the EBAPS, also claims to address this so the absence of a link between items 6 and 29 may be questioned. This is mediated by acknowledging that item 6 student written responses place prominent emphasis on what science can do (evaluate which studies are best) and not the idea that "everything's up in the air!" [68].

Items 3 and 11 When looking at the student written responses to EBAPS question 3 and EBAPS question 11, it would seem that an aspect akin to both pertains to views about science. Not general views about science, but specifically a view which is linked to what could be called "accuracy" of science. In question 3 responses it was observed that students treated the word "estimate" differently. Generally speaking, some students took it to mean an educated guess while others seemed to focus on its meaning being more like that of "exact". In question 11 responses a similar behavior exists where some students would view science as unambiguous and/or as having results which are sensitive to small errors. It may be said, then, that question 3 and question 11 share a common link representative of the aforementioned traits. This link could be one which exists due to student views about achieving precise results within science.

In both question 3 and question 11 students are also considering numeric calculations, as dictated by the context. As such, this common view would appear to be a belief regarding the extent to which computational results in science are sensitive to input variables. More succinctly, the link may be described as the extent to which students view (computational) science as chaotic. Here, "chaotic" refers to the tendency to have vastly different outcomes based on the variables utilized. The use of the word "computational" is employed because neither question 3 nor question 11 responses seem to hint at students reflecting on other conclusion types which can be made in science (e.g. observational results). One can then see how students who view (computational) science as being chaotic may view their answers on a science test as either right or wrong, while another consequence could be that these same students may lack any certainty in how they performed on a science test. Similarly, for computer simulations, a student who views science as chaotic may emphasize extreme difficulty in getting accurate results or even reliable estimates.

It is curious how this proposed view is related to views about the tentativeness of scientific knowledge, as well as views about metacognition. Does viewing science as being chaotic mean that a student will also be more likely to view the findings put forth in science as being highly unreliable? Does viewing science as being chaotic prevent students from further reflecting upon and/or monitoring their understanding? These types of questions would need further investigation beyond what has been done in this work.

Items 5, 11, and 12 Above it was discussed how EBAPS items 3 and 11 share a unique link, however, item 11 may all be grouped together with items 5 and 12 as well. With 11 functioning as the focal point of these three items (see Fig. 4.9), it may be the most important item in determining what links these together. From simply reading item 11 it would appear that it is assessing the ability of a student to monitor their own understanding (which fits with the theoretical EBAPS axis to which it is claimed to be associated, "nature of knowing and learning"). Interview data for item 11 supports this, as most student responses could be epistemically linked to them assessing their understanding (another theme involved views about science, discussed in the linking of items 3 and 11). In exploring the written response data of items 5 and 12 it was noted that students often claimed varied study methods "improved understanding". This emphasis was not explicitly present in student written responses to items 9 and 16 (recall that items 5, 9, 12, and 16 form a factor). It is believed, then, that items 5, 11, and 12 all represent items which are gauging the ability of a student of to assess their understanding of a topic.

It is possible that item 3 also belong to this same grouping, but the written response data did little to indicate this. However, in the previous discussion we put forth the idea that viewing science as chaotic could have an impact on their ability to monitor their understanding in topics of science. It is purely speculative, but in this sense item 3 could belong with items 5, 11, and 12.

Conceptually, we now see why item 11 was originally grouped with items 5, 9, 12, and 16 (see Ch. 4 on page 83). As items 5, 9, 12, and 16 are claimed to gauge the efficacy of hard work and varied study strategies, the ability of a student to assess their understanding would go a long way in determining whether varied study strategies and/or hard work is influencing their learning.

Items 13, 15, and 17 The linking discussed below will focus on that which exists between EBAPS items 13, 15, and 17. No student written response data exists for EBAPS question 13, so subsequent speculation of the item will rely heavily on researcher experience regarding the EBAPS instrument. Fundamentally what question 13 is attempting to probe is the (unsophisticated) idea that learning may occur without the need to personally engage with the material via working practice problems or reflecting on real-life examples. Question 13 belongs to EBAPS axis two ("nature of knowing and learning"), which essentially claims to assess the extent to which a student in science relies on receiving knowledge (taking presented knowledge to be true without any reflection) as opposed to constructing and monitoring their current knowledge, as well as any new knowledge.

Given the wording of item 13, it is expected that responses to questions 15 and 17 be influenced by the value students place on making real-world connections and/or the value students place on actively engaging with the material. The lack of explicit focus regarding the real-world in question 15 indicates that the former of the two possibilities mentioned above is not what is functioning as a link here. It should be noted that the use of the word “own” within question 13 when referring to working through material does open up an additional avenue for students to express particular personal learning preferences (e.g. working in groups). A prevailing theme from the written responses to question 15 involves the students actively reflecting on working through a problem. This takes on a variety of roles, from looking at the general ideas/concepts within a problem to considering the possibility of a step-by-step algorithm needed to solve the problem. Question 17 has similar overall themes in which students tend to place value on knowing the formulas or knowing the concepts behind the formulas while (possibly) being able to interpret the results.

All three of these EBAPS questions are at least partially composed of a statement which involves understanding/learning material by working through practice problems. A common belief shared among students for these three questions is more difficult to come by. Question 15 and 17 are more heavily relied upon here to help bring forth the belief dictating this link between items. Consider in written responses to item 15 that many students took on both the view that “big ideas” lacked specific information and/or the view that knowing the algorithm means getting the right answer (as these pertain to working problems). This view was offset by students who claimed that understanding the “big ideas” meant uncovering greater detail and possibly being able to solve most any similar problem scenario. In written responses to question 17 students displayed the view that you need to know when, where, and how to implement an equation to solve a problem (i.e. the algorithm to solve the problem). This view in question 17 was contrasted by a view among students which emphasized a need to know why the equation works (i.e. need to know concepts/“big idea”). It is proposed that the link connecting items 13, 15, and 17 pertains to the extent to which students view problems in physics/chemistry as being algorithmic. From the alternative responses to question 15 and 17 (discussed above) it would seem that this view can be offset by a view which emphasizes the importance of concepts, leading to the student idea that problems do not need to follow a set path/approach but rather a set of principles.

In summary, the proposition is that items 13, 15, and 17 are all linked by the extent to which students believe that working problems (likely numeric problems) in science are algorithmic, a view which is contrasted by the extent to which students believe these problems are governed by a set of principles. This view could more succinctly be described as the problem solving sophistication of the students. In theory, one would suspect that this belief be heavily influenced by the belief behind the EBAPS axis "structure of scientific knowledge", as finding concepts and connections within the material are a vital part of the proposed belief.

Items 12 and 13 Again, recall that EBAPS question 13 has no student written response data, hence all discussion of item 13 is dependent on researcher interpretation. The link between items 12 and 13 is negative, revealing that sophisticated answers for one item correlates to unsophisticated responses on the other. Furthermore, both items claim to belong to "nature of knowing and learning", axis two of the EBAPS. One may start by considering what belief would cause students to, say, value relating material to their own ideas while also believing that really clear lectures meant good students will not need to do much practice? As item 12 has been linked to Factor 1, *source of ability to learn*, we may also wish to consider what belief causes students to value hard work and good study strategies while also believing that really clear lectures meant good students will not need to do much practice.

One proposition for the construct linking items 12 and 13 is their dependence on authority vs. self, where students are either passive receivers of knowledge or active creators of knowledge. One can imagine how students who are passive receivers of knowledge believe it is possible to learn just through "really clear" lectures (unsophisticated). However, why would these same students believe that the material could be understood better if it is related to their own ideas? We would expect that passive learners do not value putting in effort needed to make personal connections with the material. Conversely, why would students who value constructing their own knowledge simultaneously not value making personal connections with science? Some insight can be gleaned from an aspect of written responses to item 12, where students express a clear concern about "translating" information. That is, some students were okay with relating material to their own ideas if the translations were accurate. It follows that in item 13 (the subsequent EBAPS item) a "really clear" lecture would ensure/display an accurate translation for their own ideas and thus it is possible to learn the material without working many problems.

Another possibility governing the link between items 12 and 13 pertains to the construct that items 13, 15, and 17 were said to measure, problem solving sophistication. To elaborate, students who do not emphasize the importance of concepts (but instead view working problems as algorithmic) may not value relating material to their own ideas (unsophisticated response to item 12) but do put emphasis on the need to work through problems so as to uncover the algorithm (sophisticated response to item 13, but for the wrong reasoning). Similarly, students who emphasize the importance of concepts may value relating material to their own ideas (sophisticated response to item 12) but not place as much importance on the need to work through problems to do so (unsophisticated response to item 13, but with potentially acceptable reasoning). Similarly, to relate this idea to *source of ability to learn*, we can see how students who value concepts may believe that "really clear" lectures will best portray the concepts at play (responding unsophisticatedly to item 13) while yet valuing hard work and varied study strategies (responding sophisticatedly to item 12).

Without written response data to item 13, it is difficult to uncover a simple construct relating it uniquely with item 12. There are too many ambiguities within the question statement of item 13 to do so, particularly in what "really clear lectures", "good students", and "lots of sample questions [...]" could mean to the students. The two possibilities put forth above should be viewed as mostly conjecture, with not enough evidence to support either.

Items 1, 12, and 26 The link discussed below will focus on that which exists between EBAPS items 1, 12, and 26 (a possible fifth-factor called *science to self*, as discussed above). In analyzing written responses to EBAPS questions 1, 12, and 26 a common theme may be found which is most closely associated with EBAPS axis two, the "nature of knowing and learning". In general, this theme is related to the value students place on integrating new knowledge with prior experiences.

Within written responses to EBAPS question 1, there was a trend among students which exhibited the belief that personal experience could/should be used as a study/learning strategy. Students here were seldom specific about why it is a valued study/learning strategy. Written responses to EBAPS question 12 has a set of students who have expressed the belief that relating material to one's own ideas improves understanding and/or makes understanding easier, in general. Similarly, there also exists a response category where students believe relating material to their own ideas is valuable because it would make material easier to remember and/or quicker to learn. Student written responses to question 26 had a very similar category in which they claimed that putting things into your own words is valuable to learning, again stating that it made material easier to memorize and/or understand.

When taking all of the evidence into consideration, it is proposed that the link between EBAPS items 1, 12, and 26 is a belief which reflects that value that students place on relating science to themselves, as it pertains to understanding new material in science. In theory, this belief would appear to rely heavily on metacognitive ability and on whether or not they should be active constructors of knowledge. It is for these reasons that EBAPS items 1, 12, and 26 have been claimed above as being strongly influenced by the EBAPS axis "nature of knowing and learning" (which also happens to be the axis to which the EBAPS authors claim these items exist). It is likely that there are also views about the real-life applicability of science involved as well, due to the explicit need to connect science in the classroom to personal experiences (from presumably outside the classroom).

Items 4, 9, and 16 The link discussed below will focus on that existing between EBAPS items 4, 9, and 16, which will be quite similar in nature to a possible fifth factor called *learning rate* (discussed above).

When assessing the written responses to EBAPS question 4, the overarching theme present was student views about natural ability and its role in learning/understanding science. Typically, natural ability would be weighed against other factors such as hard work, interest, and/or learning method/s. Written response data to EBAPS question 9 also has students considering the role of natural ability by weighing it against alternative learning methods, hard work, and/or a commitment to learning. Within written responses to question 16 it is seen that students discuss personal interest and a view that some students learn the material faster than others do, both of which were also present within question 4 responses. Both the format of EBAPS question 4 and EBAPS question 16 contain a phrase involving the speed at which learning occurs. In question 4 this shows up in the phrase " [...] most students either learn things quickly, or not at all", whereas in question 16 this appears with "given enough time [. . .]". Question 9 does not appear to explicitly allude to the speed at which learning occurs, but "natural ability" does appear within the problem statement and there exists a great deal of consideration of natural ability within the written responses.

The last piece of information to consider is that a prominent trend in the written student definitions of "natural ability" is the ability to learn material quickly. Given the evidence, it is proposed that EBAPS items 4, 9, and 16 are linked by student views regarding natural ability, specifically the belief that learning occurs at different rates for different individuals.

Items 2, 5, and 8 Although no student written response data exists for EBAPS questions 2, the simplicity of the question and the researcher's experience with all other written responses enables a discussion of the key components within the question statement. Given such, it is believed more likely than not that students focus on the word "understanding" and the phrase "remembering facts" in the context of physics and chemistry for EBAPS question 2. The intention of question 2 is to bring forth student beliefs that, since scientific knowledge is well-connected, they should be able to create more complex knowledge from more simplistic knowledge (hence the creators propose that "strongly agree" is the most sophisticated answer). However, a vast majority of students are likely to value "facts" potentially due to their classroom experience (expectations). Most traditional physics courses do treat knowledge as being discrete "facts" that are demanded to be known by the students (formulas, numeric constants, etc.). Furthermore, most traditional physics classes will also test student "understanding" through the use of numeric problems, which only further strengthens the value that students place on the need to remember these "facts".

Within written responses to question 5 there exists a subset of students who put heavy emphasis on memorization, their ability to retain knowledge about the material. Further, students who value memorization in their responses seem to imply "understanding" as being equivalent to remembering semi-discrete bits of information. It is also worth noting that question 5 is the only item on the EBAPS to contain a form of the word "study" within the question statement. It is believed likely that the word "study" invokes ideas regarding the memorization of "facts" in preparation for a test of "understanding".

In written responses to EBAPS question 8 exists a belief among a subset of students that information is fundamental, often showing up as an emphasis on theories as requiring evidence. The predominant trait relating these three questions appears to be the value that students place on knowing "facts". It is not entirely clear whether or not these same students believe connections between "facts" exist and/or actively seek them out. However, written response data from question 8 does display a subcategory of sophisticated responses which believe there to be connections within data, which is complemented by the aforementioned subcategory of unsophisticated responses in which students put an emphasis on gathering "facts" for evidence in theories.

The trait connecting items 2, 5, and 8 on the EBAPS is related primarily to either axis one ("structure of scientific knowledge") or axis two ("nature of knowing and learning") of the EBAPS. While axis one as described by the authors of the EBAPS may perfectly express the trait present, it can only be described in these findings as the importance that students place on knowing discrete information in the context of supporting understanding within science. Given the presence of item 8 among these three items, there is more evidence to suggest that this view is a reflection of students' beliefs about knowledge in science than it is a reflection of classroom expectations. Either way, this view appears to influence how students view knowing and learning within science.

Items 19 and 27 No written response data exists for EBAPS question 19, so researcher experience is relied upon. The presentation of item 19 is clear-cut, does science follow a set of rules or not? As presented, this item would seem to measure the same thing that has been claimed by items 3 and 11 (the extent to which science is viewed as being chaotic). However, items 3 and 11 pertain moreso to achieving accurate results when performing calculations in science. Item 19 differs from this in that it associates more strongly to treating science as being akin to nature and the complications that are involved in studying nature (and not in performing calculations). It is proposed that item 19 link to items 27, 28, and 30 because of how uncertainties play into addressing missing information within science (recall items 27, 28, and 30 are believed to reflect the factor *evaluating uncertainty*). However, items 19 and 27 are strongly linked not only by evaluating uncertainty but perhaps to an even greater extent by how applicable science is to real-life/nature. Both items 19 and 27 involve explicit reflection on the applicability of science to the real-world and the complications that are present in doing so.

Confirmatory Factor Analysis

Introduction

Confirmatory factor analysis (CFA) is a type of structural equation modeling designed to test the validity of a theoretical factor structure [8]. While exploratory factor analysis is a methodology driven by data, CFA is a method driven by model (theory). CFA is model-dependent in that the researcher defines which items belong to which factors, as well as any error and/or factor covariances. The following work with CFA will utilize the factor models acquired by the cumulative work put forth in Ch. 4 on page 83 and Ch. 4 on page 115.

The factor models to be tested are a four-factor model consisting of factors 1-4 as presented in Table 4.13, and a five-factor model including an additional *science to self* factor, also shown in Table 4.13.

Factor Model Covariances

Confirmatory factor analysis allows the researcher to link any item measurement variance present within the model. As such, several items believed to be sharing a common source of measurement error that is not directly accounted for by the factors of the model are proposed hereafter.

Item 9 and item 16 both belong to the factor *Source of Ability to Learn.* However, student written responses (Appendix F) along with partial correlation mapping (Ch. 4) have indicated that these items are also influenced by student views regarding natural ability. As student definitions of natural ability within their written responses vary, the precise epistemic nature of this relationship is uncertain. In general, these definitions of natural ability manifest in an overall belief that people learn material at different rates. As the *Hard Work vs. Natural Ability* factor may be responsible

Table 4.13: Descriptions for the Factors to be used within Confirmatory Factor Analysis

1.	Source of Ability to Learn (Items: 5, 9, 12, 16) The efficacy of hard work and varied study strategies. Do students believe they can learn science by devoting time and effort toward engaging with the material?
2.	Hard Work vs. Natural Ability (Items: 21, 22, 25) The influence of hard work as compared to that of natural ability, when determining the success that an individual may achieve. Here, success is generalized and may pertain to life, learning science, and doing science.
3.	Structure of Science (Items: 6, 8, 10, 14) Views about the tenets of science and the existence of science as an authoritative figure. These views involve the philosophical, sociological, and psychological aspects of the nature of science [142]
4.	Evaluating Uncertainty (Items: 27, 28, 30) How uncertainties/ambiguities within the knowledge-gathering process are reconciled. Closely related to stages within epistemic reasoning and reflective judgment [120, 125].
5.	Science to Self (Items: 1, 12, 26) To better learn science should course material be actively connected/expressed through personal experiences, ideas, and descriptions? Or, is it best to learn the material as it is presented?

for the belief influencing how 9 and 16 are responded to, it could be argued that this error covariance not be present.

Items 21 and 22 both belong to the factor *Hard Work vs. Natural Ability*. The reason to covary the measurement error of these two items is multi-purposed. First, their proximity on the exam and the nearly identical structure of their statements makes it plausible that students treat these as being predominantly the same. Second, what a student defines as "success" is not information which is available to the researcher. As both items 21 and 22 discuss hard work and natural ability in the context of being "successful", this may invoke a uniquely distinct influence as compared to item 25 (which presents a scenario in which an individual's "success" has been contextualized by the psuedo-students Anna and Emily).

Items 6 and 10 share a unique trait among the other items which belong to the factor *Structure of Science*. While all student responses for the items within *Structure of Science* find common ground in nature of science, items 6 and 10 have some students who believe science provides definitive answers. These students represent a kind of faith-in-science mindset and appear to treat science as authority without elaborating on why. For these students, it cannot be said if they are basing their answers off of sophisticated NoS views, or if they merely believe that they are supposed to trust science because they are told to. For this latter possibility, the error between items 6 and 10 are covaried.

Covariance is likely also necessary between items 12 and 26 for the factor *Science to Self*. Here, items 12 and 26 pertain to students' "own words" and "own ideas" and is contrasted by item 1 which focuses on students' "own experiences". The influence of personal experience may be strong, but there does not exist enough evidence within written student responses to indicate that one's own words/ideas create a distinct difference to one's own experiences when it comes to learning science. Items 27 and

28 on the factor *Evaluating Uncertainty* are similar in that they both pertain directly to science and nature (earthquakes, dinosaurs). This context is different from that of item 30, where evaluating uncertainty takes place during a homework session. However, student response data again did not provide enough reason to justify the need to treat this as a source of common measurement error between items 27 and 28.

In summary, both the four and five-factor models to be tested in CFA will include error covariance for items 9 with 16, 21 with 22, and 6 with 10.

Data

Data here represent the other half of the data not included within exploratory factor analysis work (Ch. 4 on page 83 and Ch. 4 on page 115),that is, these data represent all the students from section two of the introductory astronomy course which took place from fall of 2012 through spring of 2017. Again, only the post-test data from the EBAPS are utilized. The following analysis is done with the "R" programming language under the "RStudio" integrated development environment and with the IBM SPSS AMOS software [4, 5].

As both the CFA software in R and AMOS require complete sets of data, list-wise deletion was employed. During this filtering process, 1066 complete data entries remained of the 1102 total from section two (36 cases removed). Descriptive statistics for the data may be found in Table 4.14, these statistics include all EBAPS items and not just those necessary for CFA.

Analysis

There exist several possible choices when it comes to normal theory estimators employed within CFA, which are used to obtain estimates of fit indices, standard errors, and parameter values. [76]. Before a choice of estimator is made, the nature

Table 4.14: Descriptive Statistics for all EBAPS Section 2 Post-Test Data After List-Wise Deletion

Question	n	mean	s.d.	median	skew	kurtosis	s.e.
Q01	1066	2.39	1.35	3.0	-0.35	-1.45	0.04
Q02	1066	1.10	1.27	1.5	0.83	-0.42	0.04
Q03	1066	2.82	1.21	3.5	-0.93	-0.37	0.04
Q04	1066	2.48	1.07	3.0	-0.37	-0.53	0.03
Q05	1066	3.14	1.12	3.0	-1.48	1.48	0.03
Q06	1066	3.05	1.28	4.0	-0.86	-0.66	0.04
Q07	1066	2.62	1.20	3.0	-0.56	-0.70	0.04
Q08	1066	2.94	1.18	3.0	-1.02	-0.02	0.04
Q09	1066	2.80	1.13	3.0	-0.83	-0.14	0.03
Q10	1066	2.55	1.14	3.0	-0.45	-0.66	0.03
Q11	1066	2.58	1.07	3.0	-0.71	-0.07	0.03
Q12	1066	2.55	1.33	3.0	-0.57	-1.14	0.04
Q13	1066	1.35	1.23	1.0	0.88	-0.75	0.04
Q14	1066	3.02	1.09	3.0	-1.02	0.25	0.03
Q15	1066	2.14	1.02	2.0	0.02	-0.58	0.03
Q16	1066	2.92	1.10	3.0	-1.05	0.46	0.03
Q17	1066	1.86	1.22	1.5	0.23	-1.22	0.04
Q18	1066	2.55	1.29	3.5	-0.48	-1.34	0.04
Q19	1066	2.39	1.46	3.0	-0.54	-1.13	0.04
Q20	1066	2.46	1.35	3.0	-0.53	-0.92	0.04
Q21	1066	3.15	0.90	3.0	-0.75	-0.21	0.03
Q22	1066	3.02	1.00	3.0	-0.71	-0.28	0.03
Q23	1066	1.65	1.26	1.0	0.52	-0.77	0.04
Q24	1066	3.06	1.34	4.0	-0.89	-0.87	0.04
Q25	1066	2.75	1.42	4.0	-0.49	-1.31	0.04
Q26	1066	3.38	1.14	4.0	-1.50	0.75	0.03
Q27	1066	3.25	1.19	4.0	-1.20	0.03	0.04
Q28	1066	2.74	1.20	3.0	-0.61	-0.63	0.04
Q29	1066	1.80	1.57	2.0	0.17	-1.35	0.05
Q30	1066	2.73	1.36	3.0	-0.69	-0.81	0.04

of the data to be used in CFA must be assessed. Due to the categorical nature of the EBAPS likert-style data, the data is intrinsically non-continuous and not capable of exhibiting multivariate normality, two of the basic assumptions of structural equation modeling [119,134]. Fortunately, there are measures of normality that can express the extent to which data violates assumptions of normality, namely, skew and kurtosis. Univariate measures of skew and kurtosis within SEM are recommended to be below 2 and 7, respectively, to provide the most accurate results possible [37, 86, 113]. As the EBAPS data utilized in CFA may be categorized as moderately non-normal (See Table 4.14 ; skew < 2, kurtosis < 7), but not severely non-normal, it is recommended that either an maximum likelihood (ML) or weighted least square mean and variance adjusted (WLSMV) estimator be used [76].

The chi-squared fit statistic within ML modeling has been shown to be sensitive to deviations from normality and, when small sample sizes are involved (≤ 250), the Tucker-Lewis Index (TLI), Comparative Fit Index (CFI), and root mean square error of approximation (RMSEA), which is an area of concern in determining model fit [38, 110,111,113]. Additionally, as normality becomes increasingly violated the chances of committing a Type I error also increase (incorrectly rejecting a null hypothesis) due to inflation of the chi-squared statistic [38, 110, 111, 113]. Given the potential sensitivity of chi-square fit statistic within ML estimators, as well as other parameter estimates and standard errors, it has been claimed that the WLSMV estimator "... provides the best option for modeling categorical or ordered data." [8, 169]. The WLSMV estimator is not available in IBM SPSS AMOS, but may be utilized the R "lavaan" package. As both the ML and WLSMV estimators may be applicable to the data, CFA using a WLSMV estimator is done in R while CFA using ML is conducted within AMOS. Each software tests both the four and five-factor models with their associated estimators as previously detailed. Although represented within the AMOS

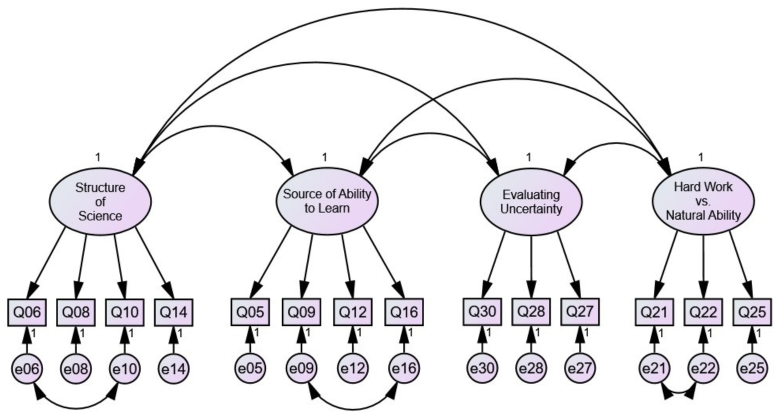

Figure 4.10: Four-Factor Theoretical Model for use in CFA

framework, Fig. 4.10 and Fig. 4.11 show the four and five-factor models put forth for CFA.

Results

Indicated previously, model fit statistics as provided via CFA are known to vary depending on the nature of the data used in testing the model, as well as the estimator. Due to this dependency there are no universally agreed upon fit statistics when it comes to determining the quality of a proposed model. As such, Table 4.15 provides the results for several fit statistics which are most frequently discussed within the literature [8, 18, 106]. Unstandardized and standardized factor loadings for CFA with WLSMV are presented in Table 4.16 for the four-factor model and Table 4.17 for the five-factor model. Similarly, the CFA WLSMV covariance estimates my be found in Table 4.18 (four-factor model) and Table 4.19 (five-factor model). The ML counterparts for these tables may be found in Appendix C.

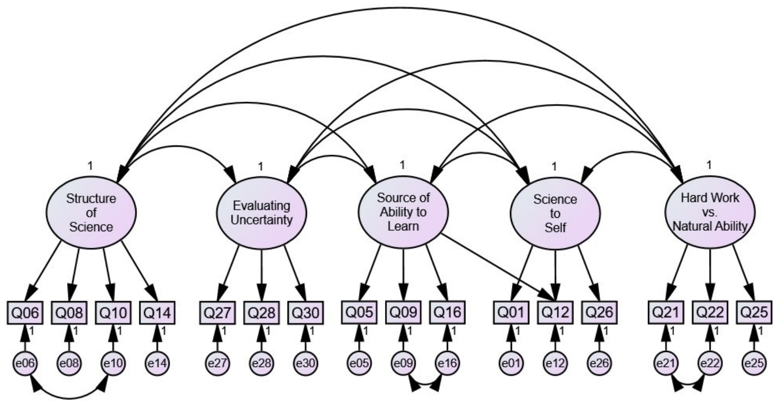

Figure 4.11: Five-Factor Theoretical Model for use in CFA

Table 4.15: Model Fit Statistics for Confirmatory Factor Analysis. df is degrees of freedom. RMSEA is root mean square of approximation. RMR is root mean square residual. GFI is goodness of fit index. CFI is comparative fit index. TLI is Tucker-Lewis index. AIC is Akaike information criterion. BIC is Bayesian Information Criterion. AIC and BIC are not available for the WLSMV method in R.
** $p < .001$

Models	χ^2	df	RMSEA	RMR	GFI	CFI	TLI	AIC	BIC
4-Factor (ML)	138.81**	68	0.031	0.040	0.982	0.968	0.958	212.8	396.8
5-Factor (ML)	196.46**	90	0.033	0.044	0.977	0.956	0.941	288.5	517.2
4-Factor (WLSMV)	125.99**	68	0.028	0.039	0.993	0.968	0.958	—	—
5-Factor (WLSMV)	180.49**	90	0.031	0.031	0.992	0.956	0.941	—	—

The model chi-squared (χ^2) quantifies the difference between predicted and observed covariance matrices, with the null hypothesis being that the proposed model covariances are equal to the covariances observed within the data. All the chi-squared values from the CFA analysis were significant at the $p < .001$ level, and thus the null hypothesis is rejected and it may be concluded that the proposed model is not a good fit for the data. However, it has already been noted that the chi-square statistic assumes multivariate normality and is sensitive to sample size. It is common for models which have good fit to still be rejected because of violations of normality [144]. Similarly, the chi-squared statistic is sensitive to sample size and will almost always reject any proposed model when large sample are used [133]. Due to these limitations, it has been proposed that the chi-squared statistic divided by the degrees of freedom be interpreted instead, with values less than 5 approaching good model fit [209]. Carmines and McIver suggest that values between 2 and 3 represent good model fit, but in general it has been observed that values between 2 and 5 represent good-fitting models [29, 138]. All models presented in Table 4.15 have values at or slightly above 2, with the exception of the 4-Factor WLSMV model, which has a value of 1.85.

The root mean square error of approximation (RMSEA) is a fit index which quantifies how well the model fits with population data when parameters are optimally chosen [111]. Literature suggests that a RMSEA values over 0.10 indicates poor model fit while values of 0.08, 0.05, and 0.01 are said to represent "mediocre", "good", and "excellent" fits, respectively [27]. All RMSEA values from the CFA analysis are found in Table 4.15, and are around 0.03.

The root mean square residual (RMR) is yet another fit statistic which quantifies the model fit by comparing the square root of the difference between residuals in the sample covariance matrix and those in the proposed theoretical model. RMR numerics are similar to those of RMSEA, where values less than 0.05 indicate good model fit while values greater than 0.08 may not be considered acceptable [25, 111]. All models put forth within the CFA have RMR values less then 0.05, typically being around 0.04, as seen in Table 4.15.

The goodness of fit index (GFI) quantifies how well the proposed model replicates the observed covariances matrix by assessing the variance and covariance present within [117,196]. This particular statistic has been shown to also have numerous issues with sample size and the number of degrees of freedom within a model [26, 188, 189]. Given the issues this index suffers from, it may be best to utilized strict criteria, expecting models to exhibit values of 0.95 or greater [189]. All CFA models in Table 4.15 have GFI values greater than 0.95, with none lower than 0.977.

Both the Tucker-Lewis index (TLI) and the comparative fit index (CFI) are measures which quantify comparisons of the proposed model to a null model (a model which would provide the worst possible fit) [132,133]. The null hypothesis within these models is that all variables are uncorrelated. A conservative bet regarding the values of both the TLI and CFI fit indices would be that values over 0.95 are representative of proposed models which are notably different than a model of worst fit [18, 111]. The four-factor model is seen to obey this criteria for model fit, while the five-factor model has TLI values which are slightly below that which is recommended (see Table 4.15).

AIC and BIC statistics are employed in comparing two models which contain the same data [14, 66]. The AIC and BIC are put forth purely for models (developed by other individuals) which may wish to compare/contrast with the models utilized within this work. These particular values are dependent on the chi-squared statistic and have parameters which inflate AIC/BIC values for more complex models. In general, the lower the AIC/BIC values, the better. AIC and BIC values should not be compared between the four and five-factor model, as these represent two models utilizing different data.

Analyzing the results from the unstandardized loadings for both the four-factor and five-factor WLSMV models (Table 4.16 and Table 4.17, respectively) indicates all items have loadings which are significantly different from 0, with $p < 0.001$. Specifically, this is evidence that these items belong to the factors that the theoretical model/s have claimed them to. The standardized loadings for the four-factor model (Table 4.16) all have values greater than 0.32, indicating that a respectable amount of variance within the item is accounted for by the theoretical factor to which it belongs [196]. Within the five-factor model (Table 4.16) all items also have loadings greater than 0.32, with the exception of item 26 (loads with 0.315 along *Science to Self*) and item 12 (loads with 0.248 along *Source of Ability to Learn*). It is plausible that Dave, within the question 26 statement, frames their response such that a non-negligible number of students now also consider the influence of authority as contrasted by the value that they place on making personal connections with science. The low cross-loading for factor 12 onto *Source of Ability to Learn* is a good indicator that its variance is indeed better explained by the factor *Science to Self*, but is otherwise not concerning as it is not a primary loading.

The proposed covariance between items 9 and 16 fails to achieve significance within both the four and five-factor models (Tables 4.18 and 4.19, respectively), indicating that there may not be a source of unique error prominent between the items, contradicting the theoretical backing for the link. The covariance placed between items 21 and 22 within the models is found to be significant at the $p < .001$ level for both models. Similarly, the covariance proposed between items 6 and 10 was also found to be significant for both models, at the $p < .01$ level.

Table 4.16: Unstandardized and Standardized Factor Loading Estimates for the 4-Factor Model with WLSMV Estimator
*** $p < .001$

Question	Factor	Unstd. Est.	Std. Est.
Q14	Structure of Science	0.704***	0.649
Q10	Structure of Science	0.387***	0.340
Q08	Structure of Science	0.731***	0.621
Q06	Structure of Science	0.520***	0.407
Q25	Hard Work vs. Natural Ability	0.788***	0.554
Q22	Hard Work vs. Natural Ability	0.564***	0.563
Q21	Hard Work vs. Natural Ability	0.392***	0.434
Q16	Source of Ability to Learn	0.737***	0.669
Q12	Source of Ability to Learn	0.664***	0.498
Q09	Source of Ability to Learn	0.748***	0.660
Q05	Source of Ability to Learn	0.661***	0.592
Q27	Evaluating Uncertainty	0.542***	0.454
Q28	Evaluating Uncertainty	0.580***	0.485
Q30	Evaluating Uncertainty	0.605***	0.444

Table 4.17: Unstandardized and Standardized Factor Loading Estimates for the 5-Factor Model with WLSMV Estimator
*** $p < .001$

Question	Factor	Unstd. Est.	Std. Est.
Q14	Structure of Science	0.702***	0.647
Q10	Structure of Science	0.392***	0.345
Q08	Structure of Science	0.738***	0.627
Q06	Structure of Science	0.509***	0.399
Q25	Hard Work vs. Natural Ability	0.792***	0.556
Q22	Hard Work vs. Natural Ability	0.562***	0.561
Q21	Hard Work vs. Natural Ability	0.389***	0.431
Q16	Source of Ability to Learn	0.774***	0.703
Q12	Source of Ability to Learn	0.331***	0.248
Q09	Source of Ability to Learn	0.795***	0.701
Q05	Source of Ability to Learn	0.676***	0.606
Q27	Evaluating Uncertainty	0.537***	0.450
Q28	Evaluating Uncertainty	0.583***	0.487
Q30	Evaluating Uncertainty	0.608***	0.446
Q01	Science to Self	0.563***	0.416
Q12	Science to Self	0.417***	0.417
Q26	Science to Self	0.315***	0.315

Table 4.18: Covariance Estimates for the 4-Factor Model with WLSMV Estimator
*** $p < .001$
** $p < .01$
* $p < .05$

Construct	Construct	Cov. Est.
Structure of Science	Source of Ability to Learn	0.815***
Hard Work vs. Natural Ability	Source of Ability to Learn	0.344***
Source of Ability to Learn	Evaluating Uncertainty	0.472***
Structure of Science	Hard Work vs. Natural Ability	0.264***
Structure of Science	Evaluating Uncertainty	0.554***
Hard Work vs. Natural Ability	Evaluating Uncertainty	0.511***
e22	e21	0.192***
e16	e09	0.014
e10	e06	0.138**

Table 4.19: Covariance Estimates for the 5-Factor Model with WLSMV Estimator
*** $p < .001$
** $p < .01$
* $p < .05$

Construct	Construct	Cov. Est.
Structure of Science	Hard Work vs. Natural Ability	0.263***
Structure of Science	Evaluating Uncertainty	0.554***
Structure of Science	Source of Ability to Learn	0.758***
Structure of Science	Science to Self	0.622***
Hard Work vs. Natural Ability	Evaluating Uncertainty	0.512***
Hard Work vs. Natural Ability	Source of Ability to Learn	0.355***
Hard Work vs. Natural Ability	Science to Self	0.177*
Evaluating Uncertainty	Source of Ability to Learn	0.429***
Evaluating Uncertainty	Science to Self	0.428***
Source of Ability to Learn	Science to Self	0.445***
e22	e21	0.194***
e16	e09	-0.050
e10	e06	0.140**

Conclusions

The theoretical four-factor model is found to be a good representation of student responses to the questions within this study, passing all prominent recommendations for model fit as put forth within literature. However, the error covariance between items 9 and 16 as proposed within the four-factor model is found to be unnecessary. The five-factor model could be said to be acceptable, having passed all fit criteria with the exception of the TLI statistic (measured 0.941, accepted cutoff at $TLI > 0.950$). Item 26 within the factor *Science to Self* may also place strain on the model, given that the factor to which it belongs explains a relatively small amount of variance within the item. Similar to the four-factor model, the five-factor model has error covariance between items 9 and 16 which is found to be unnecessary.

The items belonging to the factors within these models are not claimed to be valid measurement tools for their respective factors. Instead, these items have merely been utilized to help identify and define prominent epistemic factors influencing student responses to the EBAPS.

Changing Epistemological Beliefs with Nature of Science Implementations

This section, titled "Changing Epistemological Beliefs with Nature of Science Implementations" is an adaptation of previously published material. The following is functionally the same material as that which was originally submitted and accepted to the peer-reviewed journal *Physical Review Physics Education Research* (volume 14, issue 1, pages 010110, co-author/s: Shannon D. Willoughby), with the exception of some basic formatting modifications [115].

Introduction

This article discusses our investigation regarding nature of science (NOS) implementations and epistemological beliefs within an undergraduate introductory astronomy course. The five year study consists of two years of baseline data in which no explicit use of NOS material was implemented, then three years of subsequent data in which specific NOS material was integrated into the classroom. Our original study covered two years of baseline data and one year of treatment data. Two additional years of treatment course data have revealed intriguing new insights into our students' epistemic belief structure. To monitor the evolution of belief structures across each semester we used student pre-post data on the Epistemological Beliefs About the Physical Sciences (EBAPS) assessment. The collected data were also partitioned and analyzed according to the following variables: college (Letters of Science, Business, Education, etc.), degree (BA or BS), status (freshman, sophomore, etc.), and gender (male or female). We find that the treatment course no longer undergoes significant overall epistemic deterioration after a semester of instruction. We also acquire a more detailed analysis of these findings utilizing the aforementioned variables. Most notably, we see that this intervention had a pronounced positive impact on males and on students within the college of Education, Arts & Architecture, and those with no concentration. Lastly, whether or not students believe their ability to learn science is innate or malleable did not seem to change, remaining a rigid construct with student epistemologies.

Epistemologies may be defined as a “set of views about the nature of knowledge, knowing, and learning [71].” In general, there are two fields of thought regarding epistemologies. The first stance is that these views progress in stages, such as was pioneered by Perry and further supported in other work [104, 135, 165]. Here, a student may begin with the mindset that knowledge is inherited from authority and eventually come to the position that knowledge is constructed. Alternatively, Schommer provided compelling evidence that there exist multiple dimensions of epistemologies and that an individual has a varying degree of sophistication along each dimension, with similar work to follow this philosophy [94, 172, 183]. These dimensions have included constructs such as the tentativeness of knowledge and the innateness of ability.

More recent work has been based around epistemic beliefs being context dependent as well, that student epistemolgies can be domain specific (physics, history, mathematics, etc.) and even situation specific [96]. Our research, and thus our presentation of data with this paper, has adopted these latter philosophies for which epistemological beliefs are context dependent (domain, specifically) and exist along non-orthogonal dimensions.

In early 2012 Duncan published a paper discussing student epistemological beliefs within an introductory astronomy class [60]. His findings led him to conclude that basic inclusions of material involving the nature of science (NOS) within the classroom could have a measurable impact on students epistemic belief structures. Briefly, the nature of science may be thought of as “[. . .] the values and assumptions inherent to the development of scientific knowledge [127]." Duncan's claim is of value because, across the country, when undergraduates take a course in science their belief structures tend to deteriorate [10, 60, 172]. In other words, students' views about knowledge and learning within science become less sophisticated after taking an introductory course in science. Duncan acknowledged shortcomings within his study and put a call out to others to investigate the influence of explicit NOS materials on epistemological beliefs within the classroom. Further details of the extant literature can be found in Ref. [212]. We proceed with a discussion of the methodology and any complications therein. Results within the groups of the study are subsequently discussed, as is a brief between group comparison utilizing normalized change. Lastly, a context-dependent interpretation of effect size is presented, followed by a summary of the findings in this study.

Methods

Design In the fall of 2012 we began to investigate Duncan's claim with the collection of data from a large enrollment undergraduate introductory astronomy class (Astronomy 110). The course instructor (one of the authors) is well versed in physics education research and maintained an active classroom, frequently utilizing methods such as iClickers, group activities, and exams with a collaborative group portion. The instructor as well as the active classroom environment have remained constant throughout all years of this study.

The conclusion of the study occurred in spring 2017, having collected two years of baseline (control) course data from Fall 2012 to Spring 2014 and three years of modified (treatment) course data from Fall 2014 to Spring 2017. This large volume of data allowed us the ability to effectively explore several variables. These variables include gender, college (Letters of Science, Business, Education, etc.), degree (BA or BS), and status (freshman, sophomore, etc.). We chose to explore these variables as they were readily available and because other epistemological research has placed focus here [9, 19, 55, 92, 161, 183, 184]. Furthermore, although we focus on the claims made by Duncan, we are also motivated to explore the status quo of epistemic beliefs in the context of an introductory astronomy course.

Fundamentally, we sought to answer the question "What is the state of student epistemic beliefs within our introductory astronomy course and can basic course modifications, focused on NOS, prevent decay of student epistemologies towards science?" This overarching question branched into a subset of five research questions:

1. How do students' epistemic beliefs about science change over the course of a semester, as compared to our baseline (control) course?

2. How do students' epistemic beliefs about the sciences change after some basic course modifications and how does this compare to the changes seen in the baseline course?

3. Are there differences in epistemic beliefs when considering gender and do course modifications effect students of either gender differently?

4. Considering the college, degree, and status variables individually, are students within the baseline course more or less susceptible to epistemic change?

5. How do nature of science course modifications effect the epistemologies of students within these groups (college, degree, and status) differently?

Precise information regarding material and philosophies implemented within the modified course can be found in our original paper [212]. Within the modified course nature of science material was included at least weekly. This allowed students plenty of practice with the material, so as to improve understanding of the content [15, 127].

Implementation was not lecture based, but focused on group activities, individual exploration, and personal reflection. Students were frequently asked to interact with each other regarding course content. Nature of science encompasses a wide range of topics which we narrowed down to the following: asking students to apply the scientific method, metacognitive tasks, model development, discussing the role of skepticism in scientific discovery, and finding connections to science in their daily lives. The new material was added to the course in lectures that had not previously filled the entire time allotted to each class meeting. These simple implementations were done with the hope that, if our findings were significant, instructors would not have to restructure an entire course in order to have a similar effect.

Population As the population can be largely influential in conclusions that are made within any study, we discuss demographics. This study was conducted at a medium-sized, midwestern, land-grant institution with an open enrollment policy. Student body profiles at this university for Fall 2016 are as follows: 45% female, 55% male, 62% in-state residents, 34% out-of-state residents, and 4% international students. Of all the students, 83% identified as Caucasian, with the next largest ethnicity being Hispanic or Latino at roughly 4%. The remaining 13% of the student body consists predominantly of American Indian, Asian, African American, foreign students, and students of mixed race. Over the course of this study (Fall 2012 to Spring 2017), these profiles have remained approximately constant. Although these percentages represent the university as a whole, the instructor participating in the study thinks these values to be an accurate representation of the Astronomy 110 course population. As we analyze this system of student epistemic beliefs, it is important to note that any findings discussed in this paper are a representation of the demographics above.

Instrument To measure student epistemic belief structures we relied on the epistemological beliefs about the physical sciences (EBAPS) assessment. This 30 question forced-choice, Likert-type instrument was developed and validated by Elby et al. The EBAPS measures epistemological beliefs along five nonorthogonal axes (structure of scientific knowledge, nature of knowing and learning, real-life applicablity, evolving knowledge, and source of ability to learn), as well as an overall axis. Details regarding the axes of the EBAPS can be found in A.1, while more information regarding the EBAPS as a whole can be found at its host website [68,69]. Each axis measures student epistemological sophistication on a numeric scale from zero to four. On this scale, a zero represents a novicelike view whereas a four represents an expertlike view, with a gradient of sophistication between these values.

Procedure The initial set of data in the study is the control, or the baseline course. This course represents data taken in both fall and spring semesters from Fall 2012 through Spring 2014, in which no explicit NOS material was incorporated. The second data set is that of the treatment group, what we call the modified course. The modified course represents data taken in the fall and spring semesters from Fall 2014 to Spring 2017, where focused NOS material was incorporated. Each semester numbered roughly 400 introductory astronomy students across two course sections.

The EBAPS was administered twice each semester. The pretest was given during the first week of class while the post-test was given during the final week of class. Each student received a physical copy of the EBAPS instrument and a bubble sheet scantron on which to mark their answers. They were given 20 min towards the end of class to complete the assessment.

A voluntary survey, the vast majority of students present when the assessment was given would still participate. Average student participation numbered around

350 students across two sections. After scantrons were collected, they were parsed to check for anomalies. Any scantrons that were unfinished were discarded, as were any that displayed obvious repetition (e.g., sheets which contained only "c" responses). Our study poses minimal risk to the student and as such has IRB exempt status. From here, pre- and post-tests were compared side by side. Students who participated in both the EBAPS pretest and post-test were matched and kept, students who did not were discarded. This helps ensure a true representation of changes in student beliefs across a semester. After this filtering process there were typically between 250 and 300 sets of matched data each semester. Student response data were then transformed into relevant EBAPS information [69]. This transformation yields a numeric value (04) for every student along each of the five axes and an overall axis. Again, these axes are described in Appenidx A.1. To gauge the evolution of student epistemological beliefs across a semester, Wilcoxin signed-rank tests were used on each of the EBAPS axes. Calculations of effect size were done using the Cohens d method with pooled standard deviations. Typical effect sizes are calculated by taking the difference in the mean of two populations and dividing by the control group standard deviation, essentially providing a signal-to-noise ratio as shown:

$$\text{Effect Size} = \frac{\bar{X}_{treatment} - \bar{X}_{control}}{s_{control}} \tag{4.1}$$

In place of the control standard deviation, we have opted to use a pooled standard deviation to help account for any noise that may have arisen in the modified group that may not have been present in our baseline:

$$s_{pooled} = \sqrt{\frac{(N_1 - 1)s_1^2 + (N_2 - 1)s_2^2}{N_1 + N_2 - 2}} \tag{4.2}$$

Within this pooled standard deviation the population sizes for each group being compared, N_1 and N_2, are taken into consideration alongside their standard deviations s_1 and s_2, respectively. Unless otherwise noted, comparisons made in this paper would have $N_1 = N_2$, as they represent matched students between pretest and post-test data. When interpreting effect size a value of 0.2 may be considered a small effect, 0.5 as medium, and 0.8 as large [40]. A thorough discussion of effect sizes is presented in works by Coe, Kirk, and Rice [39, 100, 173]. We have reason to believe that effect sizes of $|d| = 0.3$ could represent "large" effects in the context of this study, see the section "Assigning Meaning to Effect Size" within this paper for further insight.

Computations in this paper were run predominantly on IBM SPSS statistical analysis software. This software allowed for simple work with this studys numerous variables and also provided a well-known statistical template from which others may conduct similar work.

Complications

Student Background This study has not dealt with several factors which have the potential to influence initial and evolving student epistemic beliefs [100, 136]. Socioeconomic factors such as parental income, education, and occupation, or even individual factors such as the number of science courses previously taken by the student are not accounted for within this study. Student performance within a course has also been linked to instructor epistemologies, however, as the same instructor was present throughout this study we do not need to deeply consider variations in data regarding a change in authority figure [194, 208].

Fall Modified Data Ideally, pretest performance between all study semesters should not yield significant deviations from each other [10]. To test this, a one-way analysis of variance (ANOVA) was performed across all semesters within the

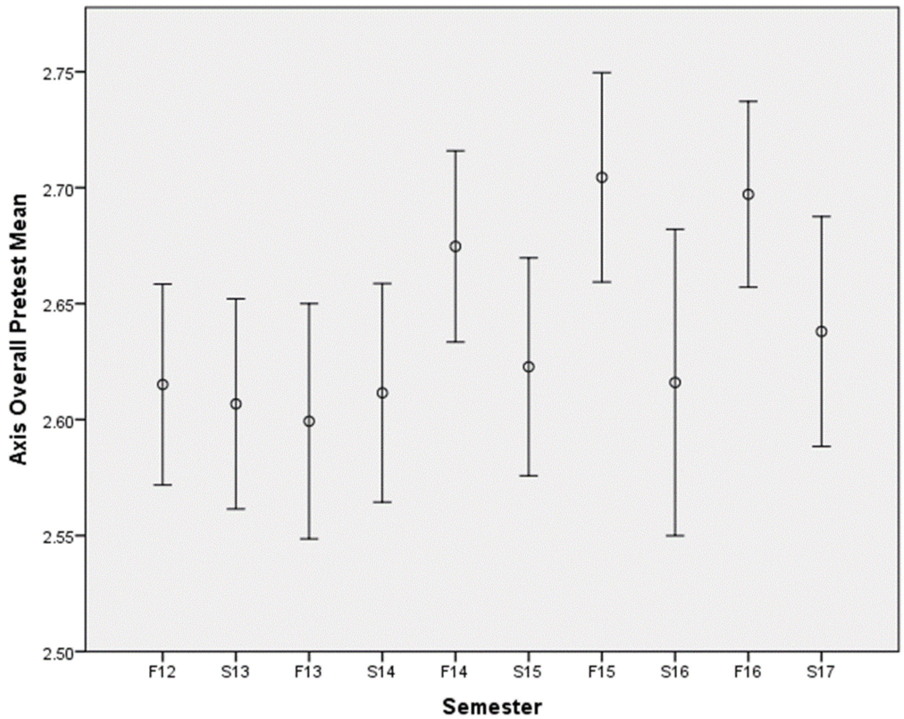

Figure 4.12: Error bar plot of average pretest EBAPS score on overall axis by semester.

study on EBAPS pretest scores (homogeneity of variances was upheld, an ANOVA assumption). Findings revealed statistically significant differences along the overall axis [$F(9, 2325) = 2.861, p = 0.002$]. Descriptives for the overall axis can be found in Appendix D.

A Duncan post hoc follow-up revealed that these significant differences were stemming predominantly from two of the fall semesters in the modified course (F15 and F16). Further probing with a Waller-Duncan post hoc Bayesian approach revealed that F14 may also be considered as contributing to this discrepancy. A plot of means by semester for the overall axis as well as the initial Duncan post hoc follow-up can be found in Fig. 4.12 and in the Appendix D, respectively. The error bars in Fig. 4.12 represent 95% confidence levels.

Note from Table D.2 that none of the spring semesters in the modified course differ significantly from those in the baseline course. The evidence seems to point towards an effect present between these modified course semesters: there are notably higher EBAPS scores in the fall as compared to the spring (and baseline). Hence, students in the fall semesters of the modified course do not begin the course epistemologically the same as those in the baseline. As best we can tell there were no noteworthy changes in the pretest environment between fall and spring semesters in the modified course, relative to the baseline course. A meeting with our Director of University Studies and the Academic Advising Center also revealed no immediate probable cause for this trend. We are open to comments from the community regarding this finding.

Based on these results, if we separately group the modified course spring semester data (S15, S16, S17) and the modified course fall semester data (F14, F15, F16), a clearer analysis forms. Figure 4.13 shows a confidence interval plot (95% confidence) of the EBAPS overall axis mean by these groupings, baseline data were also portrayed for reference. Using guidelines from Cumming and Finch, confidence intervals dictate that fall semesters within the modified course come from different populations with $p < 0.01$ [50]. Since spring students in the modified course are epistemologically similar to those in the baseline, we believe the results of the grouped data from the spring portions of the modified course (S15, S16, S17) best convey the effect our changes had on the course. As such, the grouped spring semesters of the treatment course will be presented alongside baseline results throughout this paper.

Although the fall modified students will not be considered, we wish to incorporate their data deviations into our data analysis. This is done by considering the difference in mean pretest values for the EBAPS overall axis between the fall and the spring students of the modified course, then quantifying this difference in terms

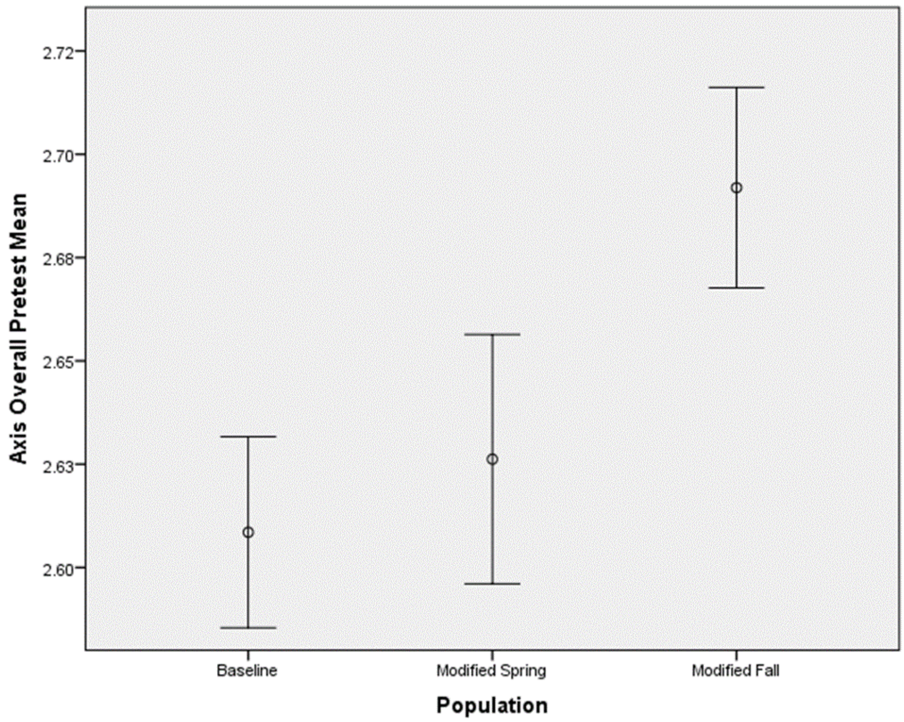

Figure 4.13: Error bar plot of average pretest EBAPS score on overall axis with data combined.

of effect size. We find that the modified fall students (F14, F15, F16) score higher on the overall EBAPS pretest axis than those in the spring with a Cohen's d effect of $d \approx 0.18$, a notably small effect and one whose cause we have yet to definitively track down [40]. As such we cannot assume that this is not just an effect due to noisy instrumentation. Thus, we move forward considering effect sizes around $|d| = 0.18$ as a discussion threshold. What we mean by this is that even if, say, Wilcoxin tests reveal significance ($p < 0.05$) along an EBAPS axis, the result will not be discussed unless their associated effect size is above $|d| = 0.18$. In truth,we will be aggressive with respect to our internal measure and use effect sizes at or above $|d| = 0.15$ with $p < 0.05$ as our significance criteria. To help emphasize the importance of effect size, we will favor the use of words such as "noticeable," "prominent," or "visible" when referring to findings exhibiting the aforementioned criteria, so as to delineate them from typical definitions of "significance" ($p < 0.05$).

Results Within Semester

Pre-post Analysis: Total populations All students within each data set were analyzed using a Wilcoxin related-samples t-test and the results are shown below in Tables 4.20 and 4.21. This analysis of the total population will yield the largest number of students for each data set within this paper and will thus yield the most statistically powerful conclusions present in this study. A bold axis label indicates prominence as previously defined ($|d| \geq 0.15$ with $p \leq 0.05$).

Baseline results: Consistent deterioration Baseline results, involving 906 matched pre-post students, concur with those from other studies in that epistemological beliefs do indeed seem to deteriorate after a semester of instruction within a science course. In particular, we find noticeable losses along axis two (metacognition) and axis five (innate vs malleable mindset). In context, this means that after a

Table 4.20: Baseline Pre-Test and Post-Test results by EBAPS axis. Scores range from 0 to 4. Bold indicates prominence with $|d| \geq 0.15$ and $p \leq .05$. Axis 1: Structure of scientific knowledge, Axis 2: Nature of knowing and learning, Axis 3: Real-life applicability, Axis 4: Evolving knowledge, Axis 5: Source of ability to learn

	Pre-Test	s.d.	Post-Test	s.d.	effect size	p value
Axis 1	2.25	0.46	2.28	0.49	0.06	0.04
Axis 2	2.60	0.52	2.52	0.56	-0.15	0.00
Axis 3	2.87	0.69	2.82	0.74	-0.06	0.23
Axis 4	2.60	0.78	2.52	0.77	-0.10	0.02
Axis 5	3.01	0.67	2.82	0.75	-0.27	0.00
Overall	2.61	0.35	2.54	0.41	-0.17	0.00

semester of typical coursework the students are less likely to engage in metacognitive processes and are more likely to believe that one cannot become better at science through increased effort or a change in study behavior.

Modified spring results: Noteworthy improvement Modified spring results (624 matched pre-post students in all) are encouraging, as overall epistemic beliefs are no longer undergoing a prominent deterioration. Axes two and five are not as pronounced as in the baseline, but do seem to be the most difficult to affect. Axis one (structure of scientific knowledge) hints at a positive response to the change, but the effect size present here does not clear our internal effect size noise filter.

Pre-post Analysis: College From now on, as data are separated, the relatively lower sample sizes will inherently lead to less statistical power, hence an increased likelihood that if significance exists it may not be detected. Fundamentally, this

Table 4.21: Modified Spring Pre-Test and Post-Test results by EBAPS axis. Scores range from 0 to 4. Bold indicates prominence with $|d| \geq .15$ and $p \leq .05$. Axis 1: Structure of scientific knowledge, Axis 2: Nature of knowing and learning, Axis 3: Real-life applicability, Axis 4: Evolving knowledge, Axis 5: Source of ability to learn

	Pre-Test	s.d.	Post-Test	s.d.	effect size	p value
Axis 1	2.28	0.49	2.33	0.49	0.11	0.01
Axis 2	2.58	0.51	2.53	0.55	-0.10	0.03
Axis 3	2.84	0.73	2.89	0.74	-0.07	0.12
Axis 4	2.64	0.82	2.61	0.77	-0.03	0.36
Axis 5	3.06	0.68	2.95	0.70	-0.16	0.00
Overall	2.63	0.38	2.61	0.42	-0.03	0.77

issue relates back to how confidence interval size is inversely related to sample size. Conclusions moving forward will consider and discuss this issue as it pertains to results.

Each particular data set may be separated by college. These colleges are Arts & Architecture (AA), Agriculture (AG), Business (BU), Education/Health & Human Development (ED), Engineering (EN), Letters & Science (LS), and University College (UC). University College represents students who have no chosen concentration and cannot be grouped to a particular common program. Other colleges were present but lacked sufficient sample sizes, most numbering fewer than 20 students total across both baseline and modified data (e.g., College of Nursing; $N = 11$), and hence were not a part of the analysis. Wilcoxin test results separated by college are shown for baseline (Table 4.22) and modified spring (Table 4.23).

Results: Positive impact for AA, ED, EN, UC students Across students of all colleges axis five (innate vs hard work) experiences prominent degradation from beginning to end of a semester within the baseline. The spring modified course is still hinting at detectable axis five decreases across essentially all colleges, BU ($p = 0.01, d = -0.24$) and LS ($p < 0.01, d = -0.36$) with certainty. In general, however, students in the modified course are now noticeably less likely to believe that innate ability dicates their learning, except those in BU and LS. Of great interest in the modified spring data is the engineering (EN) students, who not only have halted declines along axis five but now seem on the brink of positive prominence along this dimension ($p = 0.08, d = 0.21$).

Baseline AA ($p = 0.04, d = -0.23$), ED ($p = 0.03, d = -0.24$), and perhaps UC ($p = 0.22, d = -0.11$) students are most prone to experiencing a decrease along axis four (absolutism vs relativism). AA and ED modified course students no longer see noticeable decreases along axis four, yet UC students ($p = 0.08, d = -0.28$) seem unaffected and are thus still of concern. This means that AA and ED students are no longer prominently struggling with either an inability to delineate evidence-based statements from mere opinion, or the thought that all science is set in stone.

Axis three (applicability of science) remains firmly constant across all colleges in both the baseline and spring modified portions of the study. Thus, students have not undergone any prominent changes regarding their views on the applicability of science.

In the baseline course only UC students ($p = 0.02, d = -0.20$) experienced a prominent deterioration along axis two (metacognition), a deterioration which is not seen in the modified course. This finding indicates that the UC students are no longer notably worse at reflecting upon their own learning after a semester in Astronomy 110. It is worth noting there is reason to suspect that should larger sample sizes be present within each baseline college that axis two may actually be decaying noticeably (except with EN).

No college from the baseline data was seen to undergo visible change along axis one (structure of scientific knowledge). In the modified spring data ED students have undergone a noticeable increase along axis one ($p = 0.04, d = 0.19$), showing a greater ability to connect concepts within science. Outside of ED (which showed positive prominence) in the modified course, axis one seems to be on the verge of a visible increase across all other colleges except BU ($p = 0.49, d = -0.08$) and likely UC ($p = 0.33, d = 0.17$).

Overall, ED and UC students experience the greatest epistemological decay in the baseline, with AA students close behind. No colleges are observed to have prominent negative changes for the modified course. However, increased sample size for the modified course would likely see visible improvement overall for EN students ($p = 0.06, d = 0.20$), while revealing that BU students ($p = 0.09, d = -0.15$) may still undergo prominent overall negative change. With such low numbers in comparison it is difficult to state definitively but, in both baseline and modified spring, EN students appear to be least susceptible to overall noticeable negative shifts in their epistemologies, an exception being baseline axis five ($p = 0.05, d = -0.18$). Concluding this section, we postulate that changes made to the course have been of greatest benefit to the overall epistemologies of AA, ED, and UC students.

Pre-post Analysis: Degree Students within the Astronomy 110 course were pursuing primarily either a Bachelor of Arts or a Bachelor of Science degree. As these were the only degrees with a sufficient sample size, the data was parsed accordingly. Bachelor of Fine Arts was the next largest degree, with only 87 total students across the 5 years of this study. Results and sample sizes regarding BA and BS students are

Table 4.22: EBAPS Pre-Test and Post-Test results for Baseline by College. Scores range from 0 to 4. Bold indicates prominence with $|d| \geq .15$ and $p \leq .05$. Axis 1: Structure of scientific knowledge, Axis 2: Nature of knowing and learning, Axis 3: Real-life applicability, Axis 4: Evolving knowledge, Axis 5: Source of ability to learn

College	Axis	Pre-Test	s.d.	Post-Test	s.d.	effect size	p value	N
AA	Axis 1	2.33	0.41	2.36	0.49	0.06	0.34	172
	Axis 2	2.60	0.51	2.51	0.55	-0.16	0.06	
	Axis 3	2.91	0.68	2.95	0.67	0.05	0.60	
	Axis 4	2.72	0.75	2.54	0.79	-0.23	0.04	
	Axis 5	3.14	0.66	2.92	0.70	-0.33	0.00	
	Overall	2.67	0.33	2.60	0.42	-0.20	0.10	
BU	Axis 1	2.14	0.46	2.21	0.46	0.14	0.03	182
	Axis 2	2.54	0.52	2.45	0.52	-0.17	0.07	
	Axis 3	2.80	0.68	2.77	0.72	-0.05	0.72	
	Axis 4	2.47	0.79	2.50	0.75	0.03	0.66	
	Axis 5	3.00	0.66	2.78	0.76	-0.30	0.00	
	Overall	2.54	0.36	2.49	0.38	-0.12	0.10	
ED	Axis 1	2.21	0.42	2.22	0.45	0.02	0.62	113
	Axis 2	2.65	0.48	2.57	0.57	-0.14	0.09	
	Axis 3	2.76	0.60	2.65	0.74	-0.16	0.28	
	Axis 4	2.63	0.75	2.44	0.78	-0.24	0.03	
	Axis 5	2.89	0.67	2.64	0.78	-0.35	0.00	
	Overall	2.60	0.34	2.51	0.40	-0.24	0.02	
EN	Axis 1	2.31	0.52	2.40	0.44	0.17	0.37	68
	Axis 2	2.61	0.53	2.61	0.58	-0.01	0.84	
	Axis 3	2.99	0.88	2.97	0.72	-0.03	0.82	
	Axis 4	2.63	0.76	2.58	0.77	-0.06	0.85	
	Axis 5	3.07	0.72	2.94	0.83	-0.18	0.05	
	Overall	2.65	0.39	2.64	0.41	-0.02	0.55	
LS	Axis 1	2.30	0.49	2.36	0.53	0.11	0.08	202
	Axis 2	2.68	0.52	2.61	0.56	-0.13	0.12	
	Axis 3	2.98	0.69	2.92	0.78	-0.08	0.36	
	Axis 4	2.65	0.77	2.62	0.73	-0.04	0.80	
	Axis 5	3.00	0.66	2.90	0.71	-0.16	0.03	
	Overall	2.66	0.36	2.63	0.41	-0.09	0.17	
UC	Axis 1	2.22	0.44	2.17	0.46	-0.11	0.29	154
	Axis 2	2.51	0.52	2.40	0.56	-0.20	0.02	
	Axis 3	2.76	0.67	2.73	0.74	-0.05	0.75	
	Axis 4	2.50	0.82	2.42	0.76	-0.11	0.22	
	Axis 5	2.96	0.66	2.76	0.75	-0.29	0.00	
	Overall	2.54	0.33	2.44	0.39	-0.29	0.00	

Table 4.23: EBAPS Pre-Test and Post-Test results for Spring Modified by College. Scores range from 0 to 4. Bold indicates prominence with $|d| \geq .15$ and $p \leq .05$. Axis 1: Structure of scientific knowledge, Axis 2: Nature of knowing and learning, Axis 3: Real-life applicability, Axis 4: Evolving knowledge, Axis 5: Source of ability to learn

College	Axis	Pre-Test	s.d.	Post-Test	s.d.	effect size	p value	N
AA	Axis 1	2.32	0.45	2.37	0.46	0.11	0.15	92
	Axis 2	2.55	0.48	2.42	0.60	-0.24	0.07	
	Axis 3	2.88	0.73	3.03	0.66	0.21	0.13	
	Axis 4	2.64	0.74	2.66	0.80	0.03	0.99	
	Axis 5	3.03	0.66	2.95	0.73	-0.12	0.28	
	Overall	2.63	0.36	2.61	0.42	-0.05	0.89	
BU	Axis 1	2.23	0.51	2.19	0.44	-0.08	0.49	113
	Axis 2	2.55	0.48	2.52	0.55	-0.07	0.47	
	Axis 3	2.77	0.73	2.73	0.73	-0.06	0.53	
	Axis 4	2.50	0.82	2.40	0.74	-0.12	0.26	
	Axis 5	3.03	0.70	2.87	0.64	-0.24	0.01	
	Overall	2.58	0.36	2.53	0.37	-0.15	0.09	
ED	**Axis 1**	2.27	0.47	2.36	0.51	0.19	0.04	81
	Axis 2	2.65	0.47	2.63	0.54	-0.06	0.65	
	Axis 3	2.79	0.69	2.77	0.76	-0.02	0.84	
	Axis 4	2.65	0.83	2.71	0.70	0.09	0.68	
	Axis 5	3.05	0.67	2.95	0.62	-0.14	0.13	
	Overall	2.63	0.37	2.64	0.42	0.03	0.52	
EN	Axis 1	2.38	0.44	2.47	0.52	0.19	0.10	81
	Axis 2	2.63	0.56	2.65	0.55	0.04	0.76	
	Axis 3	3.02	0.73	3.12	0.69	0.14	0.35	
	Axis 4	2.62	0.78	2.68	0.72	0.08	0.60	
	Axis 5	3.06	0.77	3.21	0.68	0.21	0.08	
	Overall	2.69	0.39	2.77	0.43	0.20	0.06	
LS	Axis 1	2.30	0.48	2.36	0.50	0.12	0.14	169
	Axis 2	2.56	0.51	2.50	0.53	-0.11	0.10	
	Axis 3	2.85	0.72	2.94	0.74	0.13	0.12	
	Axis 4	2.66	0.84	2.68	0.78	0.03	0.91	
	Axis 5	3.15	0.63	2.90	0.73	-0.36	0.00	
	Overall	2.65	0.38	2.61	0.42	-0.10	0.38	
UC	Axis 1	2.22	0.59	2.31	0.47	0.17	0.33	77
	Axis 2	2.55	0.57	2.50	0.55	-0.09	0.55	
	Axis 3	2.83	0.74	2.81	0.77	-0.02	1.00	
	Axis 4	2.74	0.87	2.50	0.83	-0.28	0.08	
	Axis 5	3.03	0.72	2.95	0.77	-0.10	0.55	
	Overall	2.59	0.45	2.59	0.42	-0.01	0.92	

found below in Table 4.24 for baseline and Table 4.25 for modified spring.

Results: Improvement for BS students Baseline results indicate that BS students experience the most prominent epistemic decay, with axes two (metacognition), five (innate vs effort), and overall being the most influential. Axis four (absolutism vs relativism) may also be showing hints of trouble ($p = 0.08, d = -0.09$) while axis one (structure of scientific knowledge) appears promising ($p = 0.07, d = 0.06$). The BS modified spring students no longer experience any noteworthy overall decay such as that observed in the baseline course, but axis five losses persist. As seen in the baseline data, axis one is still promising ($p = 0.02, d = 0.12$) for this treatment data as well. To summarize, after instruction in the modified course, BS students are no longer more likely to forgo metacognitive exercises but are still noticeably more likely to believe that scientific ability is innate. Within the baseline, BA students undergo notable decreases along axis five ($p < 0.01, d = -0.26$) but no concerns otherwise. In general, the BA students in the modified course behave similarly to those in the baseline as there appears to still be issues with axis five ($p = 0.06, d = -0.22$), but no other near prominences to be concerned with. Thus, BA students remain essentially unaffected by changes made to the course.

Pre-post Analysis: Gender How male and female epistemologies begin and evolve within astronomy can be evaluated as well. Shown here are tables for baseline (Table 4.26) and modified spring (Table 4.27), separated by gender. Student gender could not be determined for as much as 15% of each set of data and thus those data were not included.

Results: Deterioration for males more pronounced Baseline males experience three prominent decay axes (two, four, and five) while males in the modified course

Table 4.24: EBAPS Baseline Pre-Test and Post-Test results by Degree. Scores range from 0 to 4. Bold indicates prominence with $|d| \geq .15$ and $p \leq .05$. Axis 1: Structure of scientific knowledge, Axis 2: Nature of knowing and learning, Axis 3: Real-life applicability, Axis 4: Evolving knowledge, Axis 5: Source of ability to learn

Degree	Axis	Pre-Test	s.d.	Post-Test	s.d.	effect size	p value	N
BA	Axis 1	2.32	0.43	2.36	0.51	0.07	0.24	205
	Axis 2	2.63	0.48	2.58	0.56	-0.09	0.43	
	Axis 3	2.96	0.68	2.94	0.71	-0.03	0.78	
	Axis 4	2.66	0.73	2.61	0.75	-0.07	0.65	
	Axis 5	3.08	0.65	2.91	0.69	-0.26	0.00	
	Overall	2.67	0.33	2.62	0.40	-0.12	0.28	
BS	Axis 1	2.22	0.46	2.25	0.47	0.06	0.07	640
	Axis 2	2.58	0.53	2.49	0.56	-0.18	0.00	
	Axis 3	2.84	0.70	2.79	0.74	-0.07	0.27	
	Axis 4	2.55	0.79	2.48	0.77	-0.09	0.08	
	Axis 5	2.98	0.67	2.79	0.77	-0.28	0.00	
	Overall	2.58	0.36	2.52	0.40	-0.18	0.00	

Table 4.25: EBAPS Modified Spring Pre-Test and Post-Test results by Degree. Scores range from 0 to 4. Bold indicates prominence with $|d| \geq .15$ and $p \leq .05$. Axis 1: Structure of scientific knowledge, Axis 2: Nature of knowing and learning, Axis 3: Real-life applicability, Axis 4: Evolving knowledge, Axis 5: Source of ability to learn

Degree	Axis	Pre-Test	s.d.	Post-Test	s.d.	effect size	p value	N
BA	Axis 1	2.30	0.45	2.34	0.49	0.07	0.28	132
	Axis 2	2.53	0.53	2.46	0.57	-0.13	0.16	
	Axis 3	2.81	0.68	2.89	0.70	0.11	0.17	
	Axis 4	2.74	0.72	2.67	0.83	-0.09	0.37	
	Axis 5	3.05	0.64	2.90	0.72	-0.22	0.06	
	Overall	2.63	0.38	2.58	0.42	-0.11	0.31	
BS	Axis 1	2.27	0.50	2.33	0.49	0.12	0.02	449
	Axis 2	2.59	0.51	2.55	0.55	-0.08	0.13	
	Axis 3	2.85	0.74	2.87	0.75	0.03	0.61	
	Axis 4	2.60	0.83	2.58	0.75	-0.03	0.50	
	Axis 5	3.07	0.69	2.97	0.69	-0.15	0.01	
	Overall	2.63	0.39	2.62	0.42	-0.02	0.98	

experience no prominent occurrences of epistemological decay. We may safely state that males are noticeably more likely to, after a semester of unmodified instruction, either value opinion over evidence or think that science is uncompromising (axis four). They are also visibly less apt both to believe that hard work does improve their ability in science (axis two) and to engage in metacognitive learning opportunities (axis two). Males no longer suffer a noticeable decay from after the modified course along these axes. Females from the baseline course undergo only prominent epistemic change along axis five, where they are more likely to view scientific ability as set in stone. This notable deterioration along axis five for females ceases to exist in the modified course. Axis two (metacognitive ability) is on the verge of troublesome for females in both the baseline ($p = 0.04, d = -0.12$) and modified ($p = 0.03, d = -0.12$) study findings.

Axis five in the modified course an excellent example of statistical power. Neither males ($|d| = 0.11, p = 0.16, N = 281$) nor females ($|d| = 0.10, p = 0.16, N = 259$) experience noticeable change along axis five as defined in this paper ($|d| \geq 0.15$ with $p < 0.05$). Yet when combined, as seen in Table 4.21, we discover a prominent change in modified spring students along axis five ($|d| = 0.16, p < 0.001, N = 624$). This also demonstrates the importance of total population results (Tables 4.20 and 4.21) in comparison to the variable-based partitioned results within this, or any, paper.

In general, baseline data reveal that males are more prone to overall epistemological deterioration than females, although females do approach prominence ($p = 0.07, d = -0.11$). After a semester of instruction in the modified course, we no longer see either gender experiencing a prominent overall epistemic decay.

Pre-post Analysis: Status The final pre-post analysis within our study considers student status: freshman (FR), sophomore (SO), junior (JR), and senior (SR). Nearly

Table 4.26: EBAPS Baseline Pre-Test and Post-Test results by Gender. Scores range from 0 to 4. Bold indicates prominence with $|d| \geq .15$ and $p \leq .05$. Axis 1: Structure of scientific knowledge, Axis 2: Nature of knowing and learning, Axis 3: Real-life applicability, Axis 4: Evolving knowledge, Axis 5: Source of ability to learn

Gender	Axis	Pre-Test	s.d.	Post-Test	s.d.	effect size	p value	N
Female	Axis 1	2.24	0.43	2.25	0.45	0.04	0.36	379
	Axis 2	2.65	0.50	2.58	0.52	-0.12	0.04	
	Axis 3	2.79	0.67	2.79	0.67	-0.01	0.92	
	Axis 4	2.63	0.75	2.57	0.74	-0.08	0.15	
	Axis 5	2.99	0.69	2.85	0.75	-0.20	0.00	
	Overall	2.61	0.34	2.57	0.37	-0.11	0.07	
Male	Axis 1	2.27	0.47	2.32	0.52	0.08	0.07	391
	Axis 2	2.57	0.53	2.48	0.57	-0.15	0.01	
	Axis 3	2.95	0.71	2.88	0.78	-0.08	0.27	
	Axis 4	2.62	0.79	2.50	0.77	-0.16	0.05	
	Axis 5	3.03	0.65	2.80	0.74	-0.33	0.00	
	Overall	2.62	0.37	2.54	0.43	-0.20	0.00	

Table 4.27: EBAPS Modified Spring Pre-Test and Post-Test results by Gender. Scores range from 0 to 4. Bold indicates prominence with $|d| \geq .15$ and $p \leq .05$. Axis 1: Structure of scientific knowledge, Axis 2: Nature of knowing and learning, Axis 3: Real-life applicability, Axis 4: Evolving knowledge, Axis 5: Source of ability to learn

Gender	Axis	Pre-Test	s.d.	Post-Test	s.d.	effect size	p value	N
Female	Axis 1	2.29	0.47	2.35	0.49	0.12	0.07	259
	Axis 2	2.63	0.50	2.56	0.54	-0.12	0.03	
	Axis 3	2.77	0.73	2.84	0.74	0.09	0.21	
	Axis 4	2.69	0.80	2.68	0.74	-0.02	0.67	
	Axis 5	3.00	0.69	2.94	0.66	-0.10	0.16	
	Overall	2.63	0.38	2.62	0.41	-0.02	0.73	
Male	Axis 1	2.30	0.51	2.35	0.49	0.10	0.09	281
	Axis 2	2.55	0.51	2.53	0.58	-0.04	0.67	
	Axis 3	2.90	0.74	2.96	0.72	0.08	0.18	
	Axis 4	2.66	0.81	2.59	0.76	-0.09	0.23	
	Axis 5	3.08	0.69	3.00	0.73	-0.11	0.16	
	Overall	2.64	0.38	2.63	0.42	-0.02	0.59	

all students could be linked with their status via online instructor-accessible campus information. The strongest cases here can be made for freshman and sophomores, as they comprised the majority of the introductory astronomy class. Results tables for baseline (Table 4.28) and modified spring (Table 4.29) are displayed.

Results: Overall benefits for freshman Baseline data show axis five (innate vs hard work) undergoing visible deterioration across all statuses with the exception of seniors, although that is a concerning axis ($p = 0.11, d = -0.14$) for them. These prominent declines along axis five ($p < 0.01, d = -0.26$) still occur for freshman in the modified spring data, but no long occur for sophomores or juniors, and seniors ($p = 0.35, d = 0.14$) are no longer at risk. This means sophomores and juniors are visibly less prone to view their scientific knowledge as set after the modified instruction, yet freshman still exhibit this.

Juniors ($p < 0.01, d = -0.35$) in the baseline were the only class to show noticeable decay for axis four (absolutism vs relativism), but ceased this decay in the modified course. One might suggest that the negative baseline change is due to a higher count of males than females for juniors who enroll in Astronomy 110. An independent Wilcoxin test revealed that JR males ($p = 0.03, d = -0.42, N = 47$) showed prominence as compared to the females ($p = 0.14, d = -0.23, N = 62$) for this baseline data. Yet with greater numbers we expect females would also exhibit significance; regardless, JR males do seem to be more prone to decay along axis four in baseline data. No other status underwent notable change with axis four in either the baseline or modified data. We thus find that juniors in the modified course no longer struggle notably, or at all, with either the idea that all scientific knowledge is relative nor the belief that this knowledge is unchanging.

As seen throughout this study, axis three (applicability of science) undergoes no visible change in either the baseline or modified portion of the study. The exception to this finding is sophomores in the modified course, who experienced prominent positive change ($p < 0.01, d = 0.25$). Hence, after a semester of instruction in the modified course, sophomores no longer have notable decay in their ability to make connections to science in their everyday lives. Axis two (metacogntion) sees prominent deterioration for freshman ($p < 0.01, d = -0.23$) and only freshman within the baseline. Therefore, these freshman are noticeably less likely to utilize metacognition after a semester in the baseline course. Unfortunately, this is an aspect of freshman epistemologies which does not change for the modified course ($p < 0.01, d = -0.18$). Sophomores remained uneffected along axis two in the modified course while both juniors ($p = 0.31, d = 0.17$) and seniors ($p < 0.28, d = 0.20$) have strong, but not significant, positive increases along this dimension.

Interestingly, axis one (structure of scientific knowledge) has a prominent increase for baseline sophomores ($p < 0.01, d = 0.15$) and is trending that way for seniors ($p = 0.09, d = 0.14$). Modified data show no significant change along axis one for any status. The significant increase for SO students along axis one as seen in the baseline is likely still present in modified spring data ($p = 0.08, d = 0.12$); however, the lower sample size (hence lower statistical power) simply is not allowing this effect to be detected. In the end, the implementations put into the modified course cause no real change compared to what was observed in the baseline for the structure of scientific knowledge.

In effectively no way do baseline seniors experience declines in overall epistemic beliefs. Meanwhile, freshman and juniors in the baseline course undergo visible overall epistemological decay and it is likely that sophomores ($p = 0.08, d = -0.13$) do as well. Although no overall epistemic deterioration is seen for any class in the modified course, it is reasonable to believe freshman still experience this ($p = 0.06, d = -0.13$).

The overarching trend in this data would seemingly be that multiple axes across all statuses other than freshman no longer undergo deterioration, and several now approach prominence with positive effect sizes (indicating improvement).

Normalized Change

To assist in comparing performance between semesters, normalized change was utilized. Normalized change is a construct put forth by Marx and Cummings and is calculated in a similar manner as the average of gains but removes students who score alike extreme values (0 or 4, in the case of the EBAPS) on both pretest and post-test [140]. The precise method of calculation is as follows:

$$c = \begin{cases} \frac{post-pre}{100\%-pre} & post > pre \\ \text{remove} & pre = post = 0\%, 100\% \\ \frac{post-pre}{pre} & post < pre \end{cases} \tag{4.3}$$

Normalized change between student pretest and post-test EBAPS scores were calculated for all axes of baseline, modified spring, and modified fall data. A Kruskal-Wallis test was then performed, comparing distributions of normalized change between the aforementioned populations. Descriptives of the populations in this comparison can be found in D.3.

Pairwise comparisons with adjusted p-values revealed that mean normalized change values along axis five ($p = 0.007$) and the overall axis ($p = 0.001$) were significantly higher for the modified spring population compared to the baseline course. Similarly, the modified spring course showed significantly higher means for normalized change on the overall axis ($p = 0.005$) in comparison to the modified fall

Table 4.28: EBAPS Baseline Pre-Test and Post-Test results by Status. Scores range from 0 to 4. Bold indicates prominence with $|d| \geq .15$ and $p \leq .05$. Axis 1: Structure of scientific knowledge, Axis 2: Nature of knowing and learning, Axis 3: Real-life applicability, Axis 4: Evolving knowledge, Axis 5: Source of ability to learn

Status	Axis	Pre-Test	s.d.	Post-Test	s.d.	effect size	p value	N
FR	Axis 1	2.18	0.43	2.19	0.44	0.01	0.91	374
	Axis 2	2.52	0.52	2.41	0.53	-0.23	0.00	
	Axis 3	2.83	0.68	2.77	0.74	-0.08	0.28	
	Axis 4	2.50	0.79	2.46	0.77	-0.05	0.57	
	Axis 5	3.05	0.66	2.86	0.73	-0.28	0.00	
	Overall	2.56	0.34	2.48	0.37	-0.22	0.00	
SO	**Axis 1**	2.28	0.47	2.35	0.49	0.15	0.00	294
	Axis 2	2.64	0.49	2.58	0.58	-0.12	0.13	
	Axis 3	2.88	0.68	2.87	0.72	-0.02	0.84	
	Axis 4	2.62	0.77	2.57	0.76	-0.07	0.33	
	Axis 5	2.98	0.70	2.78	0.76	-0.28	0.00	
	Overall	2.63	0.36	2.58	0.42	-0.13	0.08	
JR	Axis 1	2.33	0.48	2.29	0.49	-0.10	0.51	125
	Axis 2	2.63	0.53	2.61	0.52	-0.04	0.73	
	Axis 3	2.85	0.68	2.77	0.71	-0.11	0.40	
	Axis 4	2.73	0.72	2.47	0.72	-0.35	0.00	
	Axis 5	3.03	0.59	2.79	0.79	-0.34	0.00	
	Overall	2.65	0.31	2.56	0.39	-0.25	0.01	
SR	Axis 1	2.30	0.48	2.38	0.57	0.14	0.09	111
	Axis 2	2.67	0.56	2.62	0.58	-0.08	0.22	
	Axis 3	2.98	0.74	2.93	0.78	-0.06	0.83	
	Axis 4	2.73	0.79	2.65	0.81	-0.09	0.58	
	Axis 5	2.95	0.71	2.84	0.76	-0.14	0.11	
	Overall	2.66	0.41	2.64	0.47	-0.05	0.58	

Table 4.29: EBAPS Modified Spring Pre-Test and Post-Test results by Status. Scores range from 0 to 4. Bold indicates prominence with $|d| \geq .15$ and $p \leq .05$. Axis 1: Structure of scientific knowledge, Axis 2: Nature of knowing and learning, Axis 3: Real-life applicability, Axis 4: Evolving knowledge, Axis 5: Source of ability to learn

Status	Axis	Pre-Test	s.d.	Post-Test	s.d.	effect size	p value	N
FR	Axis 1	2.24	0.48	2.29	0.49	0.10	0.09	338
	Axis 2	2.56	0.49	2.46	0.55	-0.18	0.00	
	Axis 3	2.79	0.73	2.76	0.73	-0.04	0.59	
	Axis 4	2.57	0.85	2.50	0.79	-0.08	0.21	
	Axis 5	3.03	0.68	2.85	0.73	-0.26	0.00	
	Overall	2.59	0.36	2.54	0.41	-0.13	0.06	
SO	Axis 1	2.32	0.50	2.38	0.46	0.12	0.08	187
	Axis 2	2.61	0.50	2.55	0.55	-0.11	0.41	
	Axis 3	2.87	0.73	3.05	0.70	0.25	0.00	
	Axis 4	2.71	0.77	2.72	0.75	0.01	1.00	
	Axis 5	3.09	0.70	3.04	0.65	-0.07	0.24	
	Overall	2.66	0.40	2.67	0.39	0.03	0.45	
JR	Axis 1	2.40	0.45	2.41	0.49	0.04	0.70	58
	Axis 2	2.61	0.53	2.70	0.51	0.17	0.31	
	Axis 3	2.89	0.73	3.10	0.64	0.29	0.07	
	Axis 4	2.72	0.75	2.80	0.66	0.11	0.48	
	Axis 5	3.14	0.62	3.15	0.68	0.02	0.34	
	Overall	2.71	0.36	2.77	0.37	0.15	0.12	
SR	Axis 1	2.30	0.52	2.40	0.60	0.16	0.25	41
	Axis 2	2.54	0.65	2.66	0.57	0.20	0.28	
	Axis 3	3.02	0.72	2.93	0.86	-0.12	0.43	
	Axis 4	2.74	0.78	2.69	0.74	-0.07	0.71	
	Axis 5	3.08	0.74	3.11	0.65	0.05	0.61	
	Overall	2.64	0.48	2.71	0.48	0.14	0.35	

course. No other differences were found in mean normalized change between these populations. The shift in mean normalized change for modified spring data compared to baseline data is given by an effect size of $d = 0.17$ as measured along the overall axis and $d = 0.13$ as measured along axis five. Thus the changes implemented to the course have resulted in noticeably improved overall epistemic beliefs, comparatively.

Assigning Meaning to our Effect Sizes

Although Cohen provides qualifications of meaningful effect sizes, we find it more insightful to give effect sizes presented in this paper a context outside predetermined norms. We shall do this by simulating the epistemological growth of students as they progress from freshman to senior. Consider that the bulk of students at our institution often take the introductory astronomy course as one of only two required science electives within the core curriculum. Consequently, this means the vast majority of students that take Astronomy 110 do so having taken at most one previous collegiate science course. If one then considers the difference in average EBAPS pre-test scores between freshman and seniors, a glimpse of epistemic growth across a collegiate career can be acquired and represented as an effect size.

Baseline data from Table 4.28 for EBAPS overall pre-test scores of freshman and seniors lends to an effect size (using pooled standard deviation) of $d = 0.28$. While by no means definitive, an effect size of $d = 0.28$ does give us a proxy for the epistemic growth of students at our institution over their college careers as quantified by the EBAPS.

The approach outlined above may also be used for essentially any of the variables within this study, but let us present an example using the variable “college” (EN, LS, BU, etc.). In SPSS we may split the college variable up by status and acquire the overall axis baseline EBAPS pretest data of freshman for BU

($M = 2.46, s = 0.30, N = 80$), ED ($M = 2.56, s = 0.34, N = 31$), and LS ($M = 2.65, s = 0.34, N = 63$) students. Senior numbers were lower than thought acceptable, and thus were combined with juniors. We were justified in doing so, as a Mann-Whitney U test of independence showed no significant differences between SR and JR for BU ($U = 297.5, p = 0.349$), ED ($U = 186.5, p = 0.700$), or LS ($U = 451.0, p = 0.211$) with respect to the overall axis in EBAPS baseline pretest scores. Relevant data of upperclassmen by BU ($M = 2.56, s = 0.41, N = 53$), ED ($M = 2.58, s = 0.39, N = 44$), and LS ($M = 2.74, s = 0.35, N = 66$) were then subsequently acquired. It follows that effect sizes indirectly representing collegiate epistemic growth for BU ($d = 0.29$), ED ($d = 0.05$), and LS ($d = 0.26$) were calculated.

The goal in this section has been to place the effect sizes presented in this paper into context by showing what typical effect sizes may be attributed to student epistemic growth across their collegiate careers. These previous effect sizes show that while $d = 0.28$ is perhaps a robust value for this course, an effect size as low as $d = 0.05$ (like that for ED students) could also represent the cumulative effect of a collegiate career on epistemic beliefs regarding general science. With respect to both this course, the instrument being used, and the context provided, we postulate that an effect size of $d = 0.3$ may be considered a "large" effect within our findings. There is, of course, room for error within this value as issues such as attrition by major are not being accounted for.

Summary

We now summarize these findings alongside our research questions, as well as other aspects of student epistemologies within this study.

Research Questions Recall that our original motivation for this work was based around the fundamental question: "What is the state of student epistemic beliefs within our introductory astronomy course and can basic course modifications, focused on NOS, prevent decay of student epistemologies towards science?" This study was intended to supplement the sparse literature on introductory astronomy student epistemologies, as well as investigate the impact of NOS implementations. These two factors led us to address a set of research questions, the findings of which we now summarize.

1. How do students' epistemic beliefs about science change over the course of a semester, as compared to our baseline (control) course?

The baseline portion of the study yielded significant decreases in epistemic beliefs along axes two, four, five, and overall, as measured by the EBAPS instrument. These data indicate that students were less apt to engage in metacognitive practices after a semester of instruction, relying more heavily on simply absorbing information (memorization). Their also either less capable of distinguishing opinion from evidence-based argument or more likely to believe all scientific findings are set in stone. Lastly, students leaving the course were more likely to believe that scientific ability is a fixed trait and not something that can be improved with hard work. Refer to Table 4.20 for the baseline data.

2. How do students' epistemic beliefs about the sciences change after some basic course modifications and how does this compare to the changes seen in the baseline course?

When analyzing across semesters within the modified course we saw that the fall semester performed notably better on EBAPS pretest scores than their spring counterparts. The spring semesters of the modified course scored similarly to the baseline on the pretest and thus were, as measured by the EBAPS, epistemologically identical to the baseline. As such, only the spring semesters were compared to the baseline within this study. Data for the spring modified students can be seen in Table 4.21.

The modified spring students responded positively toward the NOS materials as they ceased significant deterioration of their overall belief structures and even began moving toward significant increases in how they view the structure of scientific knowledge (axis one). This means our students may now be more inclined to see science as an interconnected weaving of ideas and concepts, as opposed to a collection of isolated "facts." Declines in metacognitive ability (axis two) were not as prominent for the modified spring course, but certainly still appear to be an area of concern. The most troublesome issue after a semester of instruction is that students are still more likely to believe that their ability to learn within science is innate, as opposed to malleable.

3. Are there differences in epistemic beliefs when considering gender and do course modifications effect students of either gender differently?

Gender EBAPS information for the baseline and spring modified course can be found in Tables 4.26 and 4.27, respectively. Male students in the baseline course appear to be more prone to overall epistemological decay than females, noticeably so with metacognitive ability, evolving knowledge, and source of ability to learn (axes two, four, and five, respectively). Baseline females only experienced visible decay in that they are less likely to believe they can improve their scientific ability with hard work (axis five). After a semester of instruction in the baseline, males clearly stood apart from females in that males were more likely to either view scientific knowledge as unyielding or were less capable of distinguishing opinion from evidence-based arguments (axis four).

Modified spring data revealed no prominent deterioration for either sex along any axis, although their is good reason to believe views regarding axis five (innate vs malleable mindset) are still worrisome.

4. Considering the college, degree, and status variables individually, are students within the baseline course more or less susceptible to epistemic change?

Baseline data revealed noticeable deterioration along axis five across all college domains (Table 4.22). Arts & Architecture students as well as Education students undergo prominent decay along axis four (evolving knowledge), a trait that no other colleges exhibit. University Studies students are the only college category that experience noticeable losses in their metacognitive views (axis two).

Bachelor of Science degree pursuers experience overall epistemic decay, whereas Bachelor of Arts students do not. Both groups, however, do show noticeable losses along axis five (source of ability to learn). Of the two groups, Bachelor of Science students also experience a visible loss in their ability to engage in metacognitive processes after a semester of instruction (axis two). Table 4.24 contains the relevant information regarding these two degrees.

Within the baseline course freshman underwent significant declines on axis two (metacognition), five (innate vs effort), and overall. Views regarding axis five also underwent visible deterioration for sophomores and juniors. Sophomores were the only group to experience a noticeable increase for their beliefs on the structure of scientific knowledge (axis one) and juniors were the only class to have a prominent decay in their ability to delineate opinion from evidence (axis four). The epistemic beliefs of seniors experience no significant change. These findings regarding student status are presented in Table 4.28.

5. Do nature of science course modifications effect the epistemologies of students within these groups (college, degree, and status) differently?

Within the modified spring course (Table 4.23), the EN students represented the only college to not only cease epistemic declines along axis five but then actually move toward significant improvement here. Overall, EN students seem to have the most resilient epistemologies, experiencing minor degradations in baseline and no degradation in the modified course. None of the colleges experienced essentially any change for better or worse along axis three throughout the baseline and modified portions of the study. AA and ED students had experienced the most trouble out of any of the colleges along axis four, however, the prominent decreases along this axis ceased to exist in the modified spring course. Furthermore, ED students in the modified spring course have undergone noticeable increases along axis one.

When partitioned by degree, we no longer see any overall epistemic decay for either BA or BS students (Table 4.25). BS students also no longer undergo notable losses along axis two (metacognition), but do still seem to suffer from deteriorated beliefs regarding their views on hard work (axis five).

The overall axis for freshman in the modified spring course no longer shows visible decay, yet axes two and five remain problematic. Sophomores originally had strong decays along axis five, this behavior ceased in the modified spring course. Furthermore, in both the modified and baseline course sophomores showed improvement along axis one and even had a prominent increase along axis three in the modified course. Juniors had visible decays for axes four, five, and overall in the baseline. No significant decays were found along any axes for juniors in the modified spring results. Seniors were firm in their epistemologies, showing no prominent increase or decrease along any axes in both the baseline and modified portions of this study. It should be noted that student numbers were lower for juniors and seniors in both portions of the study. These data may be found in Table 4.29.

Fall Semesters Within Modified Course The discrepancy between spring and fall semesters within only the modified course remains a mystery. As such, we cannot be completely certain that the positive outcomes seen in the spring are attributable to solely the implemented NOS material, despite the spring population being epistemologically identical to our baseline population. Although not presented, course changes seemed to have little effect on the modified fall students. Nevertheless, what we can say is that the cause of this discrepancy does not seem to be linked to any of the variables put forth within this study. That is, the differences in the spring and fall do not seem to be caused by college, status, degree, gender, or the student proportionality within those variables between populations. Whatever the cause, students in the fall modified course scoring higher on the pretest across effectively every axis and every primary variable (college, degree, gender, status). Independent probing of the data has found that while high-school GPA and whether or not a student has taken college preparatory courses are capable of generating the differences we see, these factors cannot be definitively stated to be the cause. As we find the modified fall students to not be a true representation of baseline students, their findings are summarized briefly in Appendix D, independent of the modified spring and baseline data. Again, while the modified fall students start out epistemically more sophisticated than the baseline students, they still experience similar decay (however, the difference in post-test scores still remain essentially the same as the difference in pre-test scores between baseline and modified fall students).

Conclusions We believe the changes incorporated into the classroom, as compared to the baseline course, have the potential to create a non-negligible positive effect on student epistemologies, as seen with the epistemologically similar modified spring students. Those who benefited the most from the intervention appeared to be males. Also experiencing notable benefits were students in Education, Arts & Architecture, and those who were without a concentration (UC). Engineering students also deserve mention, as they no longer undergo negative change along any dimension and are now on the cusp of significant positive improvement for axes one (structure of scientific knowledge), three (real-life applicability), five (innate vs effort), and overall. In general, axis five has remained the most difficult to effect. We suspect it may have more to do with student views regarding study strategies (which was not a focus of the implemented material) than it does fixed ability; however, more work is needed to support this claim. Results from this study may be utilized to guide focused epistemological work for students of a particular gender, status, or college with respect to the EBAPS axes. The instructor for the course wishes to continue the use of the additional NOS material, as they have received favorable student feedback and finds the new material to make for a more compelling, engaging classroom.

Acknowledgments

We would like to thank the Montana Space Grant Consortium for funding, Arthur Bangert for statistical consultation and Diane Donnelly, our Director of University Studies and the Academic Advising Center, for her insight.

End of previously published material as adapted from [115].

Going Beyond: Item Analysis of the EBAPS

The following work was conducted after the above was published and thus does not represent peer-reviewed work. It is included here because it maintains continuity with the findings presented above. As stated, it was found that the fall students within the modified portion of the study answered more sophisticatedly on the EBAPS pretest than did those in the baseline and modified spring portions of the study. This work quantifies which items were being answered differently by fall students on the EBAPS pretest, as compared to the baseline and modified spring students. This will be followed up with a similar analysis on the post-test data. Using the baseline course as a reference, an exploration of how the material affected item responses within the modified fall and modified spring students will then be presented. In summary, we seek to expand upon the findings above by addressing the following questions:

1. How do responses to EBAPS items on the pretest differ (epistemically) between the baseline, modified fall, and modified spring students?

2. How do responses to EBAPS items on the post-test differ (epistemically) between the baseline, modified fall, and modified spring students?

3. Using the baseline course as reference, how did the material implemented in modified portion of the class affect the manner in which students responded (epistemically) to the EBAPS items?

These findings are further explored utilizing both the written response data and factor analytical results presented elsewhere (found in Appendix F and Ch. 4 on page 115, respectively).

Data Preparation In data preparation it was decided that the student responses would be treated as either a 0 (unsophisticated), 1 (neutral), or 2 (sophisticated).

Recall that the authors of the EBAPS have assigned values to every response choice for an item, values which range from 0-4 (least sophisticated to most sophisticated). As such, responses with values less than 2 were assigned a 0 (unsophisticated), exactly 2 were assigned a 1 (neutral), and greater than 2 were assigned a 2 (sophisticated). Treating response types as tripartite allows us to more easily identify changes to item responses when making comparisons between baseline, modified spring, and modified fall courses.

In order to help quantify how students may have answered differently on the individual EBAPS questions, a chi-squared test of homogeneity was utilized. The general null hypotheses being tested are:

The proportion of students answering in a novice manner is the same for baseline, modified spring, and modified fall data.

The proportion of students answering in a middling (neutral) manner is the same for baseline, modified spring, and modified fall data.

The proportion of students answering in an expert manner is the same for baseline, modified spring, and modified fall data.

We will first test these hypotheses on each EBAPS item for the pretest data, then on the post-test data. Our independent variables within the chi-squared analyses are thus the course types (baseline, modified spring, and modified fall) while the dependent variables are the response types (0-unsophisticated, 1-neutral, and 2-sophisticated) for a given question. When the chi-squared statistic is found to be significant, a z-test of two proportions post-hoc test was employed, for which Bonferroni adjustments were made when determining significance. Where the chi-squared test will reveal that at least one of the null hypotheses were false, the post-hoc test will identify which independent variables are responsible. Significance for both the chi-squared statistic and the z-test (prior to Bonferroni adjustments) was chosen to be $p < 0.05$.

Pretest Results Chi-squared tests of homogeneity revealed that at least one of the null hypotheses should be rejected for EBAPS items 2, 5, 6, 8, 9, 11, 16, 17, 22, 24, 25, and 28 on the pretest data. Table 4.30 contains the results of these chi-squared tests, while Tables D.9 through Table D.21 in Appendix D contain the post-hoc tests conducted when chi-squared values were found to be significant. Last, Table 4.31 contains a summary of the findings. When comparing significant results between data sets, the use of the terminology "more sophisticated" reflects the finding that one data set behaved as follows: contained more sophisticated responses to a question, contained fewer unsophisticated responses, or contained both more sophisticated responses and less unsophisticated responses. The use of the terminology "less sophisticated" would refer to the opposite finding/s (e.g. data set being referred to contained fewer sophisticated responses). Keep in mind that "more"/"less" refers to the percentage of students answering in a particular manner and not the raw number of students.

The baseline responses were significantly different from the modified fall responses on items 5, 8, 9, 11, 16, 22, 25, and 28. All these items, except for item 11, were found to be answered in a significantly more sophisticated manner by the modified fall students. Item 11 is peculiar because there are significantly fewer neutral responses than the modified fall students, but no other detectable differences (i.e. baseline students neither had significantly more/less unsophisticated or sophisticated responses than did the modified fall students).

The baseline responses were significantly different from the modified spring responses on items 6, 11, 17, and 24. Item 11 had baseline students answering in a more sophisticated manner, while items 6 and 17 had baseline students answering

Table 4.30: Results of chi-square test of homogeneity between Baseline, Modified Spring, and Modified Fall data on EBAPS pretest items. N = 2334

	Q01	Q02	Q03	Q04	Q05	Q06	Q07	Q08	Q09	Q10
χ^2	0.680	9.903	7.505	9.460	11.347	10.084	1.447	16.095	24.288	2.376
p-value	0.712	0.007	0.111	0.051	0.023	0.039	0.836	$p < 0.001$	$p < 0.001$	0.667
df	2	2	4	4	4	4	4	2	4	4
	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20
χ^2	18.779	2.044	0.361	4.463	7.447	16.723	15.275	4.713	3.317	6.930
p-value	0.001	0.360	0.835	0.347	0.114	0.002	$p < 0.001$	0.095	0.506	0.140
df	4	2	2	4	4	4	2	2	4	4
	Q21	Q22	Q23	Q24	Q25	Q26	Q27	Q28	Q29	Q30
χ^2	9.155	22.884	7.834	10.874	33.573	3.138	6.135	13.138	1.114	3.425
p-value	0.057	$p < 0.001$	0.098	0.028	$p < 0.001$	0.535	0.189	0.011	0.892	0.489
df	4	4	4	4	4	4	4	4	4	4

in a less sophisticated manner. Item 24 was another oddity in that there were more neutral type responses seen in the modified spring students, but no other notable differences between sophisticated or unsophisticated responses.

Modified spring and modified fall responses differed on items 2, 8, 9, 16, 22, 25, and 28. Modified fall students were more sophisticated on all the aforementioned items except item 2, for which the modified spring students were more sophisticated.

In comparison to the modified fall students, both baseline and modified spring students had less sophisticated responses to items 8, 9, 16, 22, 25, and 28. Modified spring and modified fall students were both notably different from the baseline for item 11. There were no unique items in which both the baseline and modified fall students which were significantly different in response type to the modified spring students.

Table 4.31: EBAPS pretest items within the chi-squared tests which were significantly different between data sets. $p < 0.05$

Data Set Comparison	Items
Baseline and Modified Fall	5, 8, 9, 11, 16, 22, 25, 28
Baseline and Modified Spring	6, 11, 17, 24
Modified Spring and Modified Fall	2, 8, 9, 16, 22, 25, 28

Post-test Results Chi-squared tests of homogeneity revealed that at least one of the null hypotheses should be rejected for EBAPS items 3, 6, 8, 9, 14, 17, 21, 22, 24, 25, and 30 on the post-test data. Table 4.32 contains the results of these chi-squared tests, while Tables D.22 through Table D.32 in Appendix D contain the post-hoc tests conducted when chi-squared values were found to be significant. Last, Table 4.33 contains a summary of the findings.

The modified spring responses were significantly different from the baseline responses on items 6, 8, 14, 17, 21, 22, and 25. All these items were found to be answered in a significantly more sophisticated manner by the modified spring students.

The modified fall responses were significantly different from the baseline responses on items 3, 6, 9, 14, 17, 21, 22, 24, 25, and 30. All these items were found to be answered in a more sophisticated manner by the modified fall students. Item 8 could also likely be grouped in with these other items, despite confirmation by the z-test (see Table D.24).

Between the modified fall and modified spring students, items 9, 24, and 30 were answered in a significantly different manner on the post-test, all of which were found to also be answered in a more sophisticated manner by the modified fall students.

Table 4.32: Results of chi-square test of homogeneity between Baseline, Modified Spring, and Modified Fall data on EBAPS post-test items. N = 2334

	Q01	Q02	Q03	Q04	Q05	Q06	Q07	Q08	Q09	Q10
χ^2	1.288	2.985	12.533	6.330	6.253	14.892	2.543	8.233	19.605	1.250
p-value	0.525	0.225	0.014	0.176	0.181	0.005	0.637	0.016	0.001	0.870
df	2	2	4	4	4	4	4	2	4	4
	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20
χ^2	1.182	0.332	3.965	10.901	0.293	7.370	13.479	1.190	2.073	8.228
p-value	0.881	.847	0.138	0.028	0.990	0.116	0.001	0.552	0.722	0.084
df	4	2	2	4	4	4	2	2	4	4
	Q21	Q22	Q23	Q24	Q25	Q26	Q27	Q28	Q29	Q30
χ^2	28.840	22.751	1.140	10.598	18.963	8.074	2.482	3.832	3.206	12.607
p-value	$p < 0.001$	$p < 0.001$	0.888	0.031	0.001	0.089	0.648	0.429	0.524	0.013
df	4	4	4	4	4	4	4	4	4	4

To summarize by grouping, in comparison to the baseline students both modified spring and modified fall students had more sophisticated responses to items 6, 14, 17, 21, 22, and 25 (and, likely, item 8) on the post-test. Modified spring and baseline students were both notably different from the modified fall for items 9, 24, and 30. There remain no unique items in which both the baseline and modified fall students were significantly different in response type to the modified spring students.

<u>Conclusions</u> The results discussed hereafter will draw heavily on references to previous work, namely the student written response data (Appendix F) and the researcher-defined factors with factor analytical work (predominantly Ch. 4 on page 115 and Ch. 4 on page 152).

Table 4.33: EBAPS post-test items within the chi-squared tests which were significantly different between data sets. $p < 0.05$

Data Set Comparison	Items
Baseline and Modified Fall	3, 6, 9, 14, 17, 21, 22, 24, 25, 30
Baseline and Modified Spring	6, 8, 14, 17, 21, 22, 25
Modified Spring and Modified Fall	9, 24, 30

Pretest The pretest findings help strengthen the previous claim that the modified fall students are indeed notably different, epistemically, from both the baseline and modified spring students when entering the astronomy class. First, let us consider the items in which both the baseline and modified spring students answered differently from the modified fall, items 8, 9, 16, 22, 25, 28. For every one of these items, the modified fall students answered in a significantly more sophisticated manner compared to both the other data sets. Furthermore, all these items may be related back to a factor found within the factor analytical work. Items 9 and 16 belong to the factor "source of ability to learn", which represents the efficacy of hard work and varied study strategies. Items 22 and 25 belong to the factor "hard work vs. natural ability", which represents the influence hard work has in comparison to natural ability when achieving success within science. With these four items (9, 16, 22, and 25), there is now strong evidence that the modified fall students entered the astronomy classroom with notably more sophisticated beliefs than either the baseline or modified spring students. Specifically, the modified fall students were more apt to display the belief that you can learn science by doing science, and that hard work is more important than natural ability. The presence of item 8 (factor "structure of science") hints that the modified fall students may have also entered with more advanced views of what a “theory” is, while item 28 (factor “evaluating uncertainty”) could indicate that they were also more capable of reconciling missing information within their knowledge.

Item 11 was answered in a more sophisticated manner (by essentially both the spring and fall modified students) indicating that the baseline students may have had slightly more sophisticated ability to assess their own understanding. As this possibility is evidenced by only one item, we find it unlikely that it plays a prominent role in belief mannerisms moving forward.

The baseline and modified spring students were not perfectly identical at the beginning of the class, differing in their responses to items 6, 11, 17, and 24. Of these four items, question 24 is unreliable due not only to the influence expectations have in the student responses, but also because it attempts to measure two rather distinct epistemic beliefs. As item 11 was discussed above, this leaves items 6 and 17 for elaboration. Item 6 is a reflection of the factor “structure of science”, namely the philosophical and sociological aspects of science and scientists. Item 17 did not belong to a factor but can be seen as a tendril within Fig. 4.9. Item 17 can be thought of as representing a kind of “problem-solving sophistication” in that is gauges the importance students place on knowing the concepts behind a formula, as well as how to interpret the outcome. Item 17 also suffers, to some extent, by delving into student expectations (a construct which the EBAPS authors attempted to avoid assessing as best as possible). Again, the scarcity of items relating to a common construct indicates that the baseline students and modified spring students do indeed represent approximately epistemically identical populations. The presence of items 17 and 24 might actually be indicating that the modified spring students come in with different expectations for the class than do the baseline students. However, the EBAPS itself was designed to elucidate personal epistemic beliefs and not expectations, epistemic differences are what is valued in this study.

Post-test As the modified spring students came into astronomy nearly epistemically the same as the baseline students, the effect of the implemented course materials are best expressed through baseline and modified spring comparisons. To begin, items 11 and 24 no longer differ between these items sets. For item 11, this is attributable to a decrease in sophistication for the baseline students while the spring students remained essentially unchanged. For item 24, this was due to an increase in the number of neutral responses, with decreases in response frequencies from both the unsophisticated and sophisticated options. Unchanged from pretest findings were items 6 and 17 (both still answered in a more sophisticated manner on the post-test by modified spring students). New to the post-test were the presence of items 8, 14, 21, 22, and 25 as being answered in a more sophisticated manner by the modified spring students. The general trend for all these items is that the modified spring students were unchanged while the baseline students experienced decay, hence creating the significant differences. The resiliency of items 6, 8, and 14 indicate that students in the reformed class no longer suffer from deteriorating views regarding the *structure of science*. Similarly, the presence of items 21, 22, and 25 indicated that students no longer experience as prominent of decay in *hard work vs. natural ability*.

In comparing the modified fall with the baseline, we see that response profiles for items 5, 11, and 28 no longer differ between these students. In the case of item 5 and 28, this appears due to modified fall students becoming less sophisticated than those in the baseline, while the reverse is true for item 11. Responses to items 3, 6, (8, it could be argued) 14, 17, 24 and 30 are now all significantly different between the baseline and modified fall students, however both still experienced epistemic decay. Item 17 was the exception to this trend, experiencing growth for both the baseline and fall students. The presence of items 6 and 14 (and 8) indicate that the included class material may have made the modified fall students more resilient to epistemic decay with respect to the factor *structure of science*. The researcher avoids making any further claims given that the modified fall students did not start out epistemically the same as the baseline students, hence making statements about the effectiveness (or lack thereof) of the implemented course material obscure.

CONCLUSIONS

This dissertation was focused around the impact that nature of science curricula had on the epistemologies of students within an introductory astronomy class. In analyzing the epistemic changes generated by the additional course material, the epistemological beliefs of astronomy students themselves had to be investigated in finer detail. This led to the presence of a complete epistemological framework describing how astronomy students were responding to the instrument being used to probe their epistemic beliefs, the epistemological beliefs assessment for physical science. These findings will help guide other researchers and instructors regarding how nature of science material can be used to impact the epistemologies of introductory astronomy students, as well as help outline key epistemic structures of students in this domain.

Introductory astronomy courses are in a unique position for these results may be extended into physics and may also be useful to a variety of other non-STEM courses. The importance of research within introductory astronomy courses arises from the presence of numerous STEM and non-STEM majors, which may be utilized to better help understand common epistemological beliefs between each. Furthermore, as many of these students are non-STEM majors taking one of the only science courses that they'll experience within their collegiate career, it is of crucial importance that their epistemic beliefs toward science be investigated [52, 84]. This is given heightened significance knowing that student epistemologies are typically seen to decay after a semester of instruction within traditional collegiate science courses [10, 54, 139, 172]. These results, then, not only detail the impact that nature of science material can have on student epistemologies in introductory astronomy, but also assist in bridging the epistemological gap between non-STEM students and students entering STEM courses such as introductory physics.

Summary

This research is centered around the epistemological beliefs of students within an introductory astronomy class. The origins for the study of student epistemologies dates back to William Perry in 1970, when he proposed a stage-like progression of intellectual and ethical development. His theory on how students think about knowledge and what it means to know extends into four basic domains for which most subsequent work on epistemologies may be framed. The first of these domains is known as "dualism", where all knowledge is treated as being either right or wrong. In this domain, knowledge that is right is held by a figure of authority while all others are "pretenders to the right answers [. . .]" [165]. The next stage,"multiplicity", is where individuals will first acknowledge that authority may not hold the right answer yet, but that a right answer does still exist. Two primary views manifest among students in this stage beyond the acknowledgment that an authority not have access to the right answer. The first is the view that, in these domains of uncertainty, students will defer to the opinion of authority. The second view is summarized by the belief that "everyone has a right to his own opinion" [165]. Students no longer defer to the opinion of the authority, they now hold all ideas about uncertain knowledge on equal footing. The third stage is defined by "relativism", in which students now hold nearly all knowledge as being relativistic. Although some answers may be more accurate than others, there now exist no right or wrong answers except in unique cases. Finally comes the stage of "commitment", where students take a position regarding topics of knowledge and knowing. These commitments are defined by frequent reasoning, evaluating, and reconsideration of their knowledge. Many of the epistemological frameworks developed since (Belenky, Magolda, Kitchener, Kuhn, etc.) may still fit into this general progression as originally proposed by Perry.

Within the study of physics the more general "epistemologies" has become "attitudes" where there is no longer just a focus on the nature of knowledge and knowing but also an inclusion of views about intelligence, learning, motivation/interest, and expectations. This treatment of epistemologies by physicists dates back to the work of Marlene Schommer in the early 1990's [183]. Schommer no longer treated student epistemological progression as being stage-like, but instead viewed beliefs as a set of more-or-less independent constructs with each having a gradient of unsophisticated to sophisticated views. Schommer's work first coined the term "epistemological beliefs" to include views about intelligence and learning as well as beliefs about knowledge and knowing. Since Schommer, physicists have chosen instead to focus less on the theoretical framework involved and more on what student "attitudes" exist and what can be done regarding these "attitudes" in the classroom. Within this current work, the term "epistemological beliefs" has been chosen, as the beliefs being studied do not include explicit consideration of motivation/interest or expectations, although they are observed to exist.

In the development of a theoretical response structure for the EBAPS, there were seen to be five prominent epistemic constructs and a set of secondary constructs. The first primary construct to be discussed has been coined *source of ability to learn* (EBAPS items 5, 9, 12, and 16), representing the efficacy of hard work and study strategies. This belief may be best represented by student responses to EBAPS question 16: "Given enough time, almost everybody could learn to think more scientifically, if they really wanted to." Another prominent construct is *hard work vs. natural ability* (EBAPS items 21, 22, and 25), which has students gauging the influence of hard work in comparison to the influence of natural ability in the context of being successful. Here, students seem to treat "success" as pertaining more towards life/career accomplishments (possibly within science) than it does towards learning. EBAPS question 22 may express this best by asking "to be successful at science . . . ", for which students must then choose from a set of options such as "natural ability and hard work are equally important." This belief is also seen to be measuring a largely independent belief, with respect to other primary and secondary beliefs. The next construct is termed *structure of science* (EBAPS items 6, 8, 10, 14), and is associated with student beliefs about the philosophical, sociological, and psychological aspects of the nature of science. EBAPS question 8 is an example of what this belief is testing: "Scientists should spend almost all their time gathering information. Worrying about theories can't really help us understand anything." Interestingly, aspects of *structure of science* are seen to be influencing student responses within many of the constructs (both primary and secondary) present within this work. A fourth construct (and another largely independent belief) called *evaluating uncertainty* (EBAPS items 27, 28, and 30) pertains to the manner in which students address uncertainties/ambiguities within the knowledge-gathering process. EBAPS question 28 outlines typical scenarios in which this belief is utilized by asking why scientists can not agree on a particular topic (in this case, the extinction of the dinosaurs). These types of questions activate modes of reasoning and judgment within students, which appear to be consistent in contexts of missing or uncertain knowledge. The fifth and final primary construct is named *science to self* (EBAPS items 1, 12, and 26) and represents the extent to which a student believes that actively making connections between course material and their personal experiences/ideas is beneficial to their learning. EBAPS question 12 succinctly summarizes the general idea behind this belief, asking for a student response to: "when learning science, people can understand the material better if they relate it to their own ideas." All five of the primary constructs were supported by confirmatory factor analytical work, written response data, and partial correlation mapping.

The primary construct *source of ability to learn* essentially resembles, epistemically, an axis of the same name on the EBAPS and has been seen within the MPEX as the "effort" axis and on the CLASS as the set of axes which make up "problem solving". *Source of ability to learn* is aligned with views about the malleability of knowledge, as outlined by Dweck, in regards to the mindset that you can become more proficient in science by working at it [62]. Also related to mindset is the axis *hard work vs. natural ability*, extending beyond hard work by explicitly comparing it to the role of natural/innate ability. This direct probing of student views of hard work versus natural ability is not measured by the other prominent epistemic instruments, although it could be argued that measuring the efficacy of hard work, which the MPEX and CLASS do as stated above, indirectly probes views about natural ability. The presence of *source of ability to learn* and *hard work vs. natural ability* would seem to contradict this, that views about the role of natural ability and the role of hard work should be probed both independently (*source of ability to learn*) and in direct opposition to one another (*hard work vs. natural ability*). Views associated with the *structure of science* are either implicit or absent from both the CLASS and the MPEX, with the exception of axes associated with the linking/connection of classroom science concepts to the "real-world" in each instrument. *Structure of science* is most strongly related to measurements made by the VNOS instrument, which quantifies views about the nature of science (views which are similar to those as presented by [142]). An interesting result with *structure of science* is the possible dependency on sophisticated views regarding NoS in order to treat science as an authority. This alludes to the possibility that at least some students must have a developed understanding of what science is (must exhibit some sophistication in nature of science) before treating it as an authoritative figure more in-line with Perry's scheme [165]. To elaborate further, some written responses indicated that students appear to treat science as a peer whose ideas are equally valid, akin to what is seen in Perry's Multiplicity stage/s. However, these responses did not allude to them as actually belonging in this stage. In the first stage of Perry's scheme, there are already established authoritative figures in academia, whereas some of these students appear to have never viewed science as an authoritative figure to begin with. As such, there may be two (or more) developmental paths within science. One path in which science is initially treated as an authoritative figure holding all the knowledge, for which Perry's scheme might apply, and another path in which students do not initially view science as an authoritative being and thus develop differently than what Perry proposes. For the latter of these two paths, it seems possible that once more sophisticated views about NoS are developed that students go from a type of multiplicity-like stage into a stage resembling that of Kuhn's Evaluatist (or Perry's Relativism/Commitment) [125]. This discussion regarding *structure of science* as it pertains to student development is nothing more than conjecture, meant to stimulate thought, as the purpose of this work has not been to propose or challenge any theories of knowledge and neither has this research been fundamentally constructed to do so. Also in-line with the findings of Kuhn was the axis *Evaluating Understanding*. This axis is largely absent within any of the aforementioned prominent epistemic instruments for science, with perhaps an exception being the "sense-making/effort" axis of the CLASS. However, it is present in the literature as functionally aligned with what Kuhn has developed and termed the "reflective judgment" model and what Perry would associate with aspects of relativism/commitment [125,165]. Lastly, *science to self* shares similarities with the "real world connection" axis of the CLASS and, by the nature of the factor, axes related to the efficacy of hard work as previously discussed. With the exception of *science to self*, all these epistemic constructs can be strongly linked to prominent epistemological components identified in literature [104].

The secondary constructs are such because they were not observed to be primary factors within exploratory factor analysis and were thus not subjected to confirmatory factor analysis. However, they are believed to still hold some merit given their presence within partial correlation mapping and the support they received within student written responses. There are six distinct secondary beliefs within this research, four of which link at least three items. The first contains EBAPS items 5, 11, and 12, for which it is believed that they are measuring an aspect of metacognition. These items are associated with student ability to assess their understanding of a topic and may be referred to *assessing understanding*. EBAPS items 13, 15, and 17 all share a common link believed to be associated with what can be called *problem solving sophistication*. At the extremes, students here would be said to treat physics problems as being either algorithmic (governed by a set methodology) or as being dependent on context (governed by concepts). The next construct may be termed *learning rate* and involves EBAPS items 4, 9, and 16. The common trend among these items is "time", albeit either implicit or explicit. Responses to these questions are believed to hold a clear dependency on the belief that students learn at different rates (but can learn the material, nevertheless). EBAPS items 2, 5, and 8 make up the next construct, referred to as *discrete knowledge*. This belief gauges the extent to which students value knowing discrete information (facts, formulas, etc.) in the context of understanding science. Although little evidence exists within the student written responses, it is believed that students who highly value knowing discrete information are opposed by students who value making conceptual connections between this discrete knowledge. The next secondary construct observed in the data is associated with the link between EBAPS items 3 and 11. This belief could be called *science as chaotic*, given that it appears to represent how students view results within science.

Specifically, this construct is outlining the extent to which computational/numeric results with science are viewed as being highly sensitive toward the initial input/s for the problem at hand. Some students may view results in science as having highly different outcomes based on minor manipulation of initial variables. This view is contrasted by students who view scientific results as behaving like more of an estimation than an precise exclamation. This student view is currently only believed to be applicable toward scientific results involving numeric calculations. Last is the construct involving EBAPS items 19 and 27, what may be called *applicability of science*. As the name suggests, this belief represents the extent to which students believe science as being capable of explaining what is observed in the real world.

Of the secondary constructs, only two are not explicitly present on either the MPEX or the CLASS (or within literature), *learning rate* and *science as chaotic*. *Science as chaotic* is almost certainly underlying aspects of "problem solving" and "sense-making/effort" on the CLASS, as this belief is strongly related to numeric calculations. As it pertains to findings on the EBAPS, *science as chaotic* may play a notable role in metacognitive activities for learning science, or as the EBAPS would call it "nature of knowing and learning." Students with unsophisticated views about *science as chaotic* may be less likely to engage in metacognition, and visa-versa. Consider an example where a student has taken an exam in physics and received their results. On questions that the student got wrong, they may not be inclined to reflect upon why they got the problem wrong because they may believe that achieving the exact answer is so difficult. This student may have attributed their incorrect answer to a slight numeric error that they assume they made, which propagated into an imprecise and incorrect result, instead of actively reflecting on the true reason/s their answer was incorrect. More research regarding how the relationship that the views embodied by *science as chaotic* have with metacognitive abilities would be beneficial to furthering understanding of student epistemic behavior in science. "Quick learning" is a prominent epistemic belief identified by Schommer in the early 1990's [183] and is an epistemic belief which *learning rate* shares some commonality with. Recall that "quick learning" itself is represented by a sophisticated extreme in which students view learning as being a gradual process, and an unsophisticated extreme in which students believe learning to occur quickly or not at all. In this work, it is unclear what the spectrum of *learning rate* represents, that is, it is sophisticated to believe that people learn at different rates but what the unsophisticated view represents is not well defined in the data. The seemingly logical counterpart to the sophisticated extreme of *learning rate* would be the belief that everybody learns at the same rate, but why does this extreme not explicitly appear? Given the association of *learning rate* to *source of ability to learn* within the partial correlation networking (and in with both EFA work and written responses), it may be that the unsophisticated extreme actually relates strongly to the efficacy of hard work and good study strategies. If the unsophisticated extreme of *learning rate* be the belief that learning occurs quickly or not at all (instead of everybody learns at the same rate), it could be that students in this extreme actually exhibit mannerisms akin to those in the unsophisticated extreme of *source of ability to learn* (they may not believe you can get better at science by doing science). What this possibility would imply is the belief that you cannot get better at science by doing science may actually be related to views about the rate at which people learn and/or views about the malleability of knowledge. This could indicate that Schommer's "quick learning" then be a blending of two constructs seen in this study as *learning rate* and *source of ability to learn*. Although this is again merely conjecture, potentially worthy of being investigated further within a study designed to do so.

Student epistemological beliefs were monitored over the course of two years using the EBAPS as a pretest and post-test measure. During this time, no changes were made to the astronomy class. In baseline results students experienced significant epistemic decay along EBAPS axes 2 ("nature of knowing and learning"), 5 ("source of ability to learn"), and overall. This implies, from axis 2, that students are less likely to engage in metacognitive behavior and/or value actively working through material and relating it to personal ideas/experience. Axis 5 indicates that these students are now also more inclined to believe that being good at science is the result of fixed natural ability. Overall epistemic decay was also seen, which is in-line with findings from physics courses across the country [10, 54, 139, 172]. When separated by gender, a clear difference exists in that males appear to be more susceptible toward epistemic decay along axis 4 ("evolving knowledge") than do females. Indicating that after a semester of instruction, males are more likely to treat scientific knowledge as being either set in stone or as completely relativistic. When separated by class standing (freshman, sophomore, etc.) many of the same changes were observed, namely significant deterioration along axes 2, 5, and overall. A notable exception to this was the performance of seniors who experienced no significant losses along any axis, including overall.

In three subsequent years, explicit nature of science material was incorporated into the astronomy classroom in such a manner that the new material was added to lectures which already had time available. The effectiveness of the additional material was obscured by the presence of alternating behavior between spring and fall students within this modified course. It was observed that the fall semesters of the modified course experienced essentially the same epistemic behavior as the baseline (control). However, the spring semesters of the modified course were notably more promising. In the spring semesters, there was no longer observed to be a significant overall decay of epistemic beliefs. There was, however, still observed to be a significant decrease in epistemic beliefs along axis 5. When separated by gender, there was observed to no longer be any significant epistemic decay. Despite the absence of significant decay, it is likely that it still exists along axis 2 for females and axis 5 for both males and females (the lower numbers due to splitting into genders made it more difficult to statistically detect the decay). Freshman exhibited similar behavior as the baseline experiencing significant epistemic decay along axes 2, 5, and overall. However, sophomores, juniors, and seniors are no longer observed to experience essentially any epistemological deterioration along any axis.

To better help understand the impact that the nature of science material had on the modified classroom, a question-by-question analysis of the EBAPS was undergone. In responses to the pretest questions, it was observed that items 6, 11, 17, and 24 were answered in a significantly different manner between the baseline and modified spring classes. However, these questions do not coherently associate with any of the epistemic constructs previously identified. Both the modified spring and the baseline courses differed significantly in response frequency to the modified fall course on EBAPS items 8, 9, 16, 22, 25, and 28, all of which may be linked to epistemic constructs previously seen. Furthermore, these items were all answered in a significantly more sophisticated manner by the modified fall students. These findings indicate that the modified fall students entered the classroom more apt to believe that they can learn science by doing science and that hard work is more important than natural ability. Item 28 also indicated that modified fall students may have been more capable of making sophisticated judgments and reflection with respect to uncertainties in knowledge. Post-test item analysis that modified spring students now have significant differences to baseline students in their response frequencies toward items 6, 8, 14, 21, 22, and 25. These differences arise due to a resiliency of the modified spring students, experience less of a decay than did the baseline students. This indicated that the included material impacted both the *structure of science* and *hard work vs. natural ability* factors. Items 6 and 14 were two new items to the modified fall students which were answered significantly differently compared to the baseline students. Again, these items represented items which were resilient with repsect to the baseline behavior. This further provides evidence towards the possibility that the implemented material effected student beliefs regarding *structure of science*. As the modified spring students epistemically best represent the baseline students (they have pretest behavior which is largely identical), the impact of the additional course material appears to have helped prevent significant deterioration of beliefs associated with the *structure of science* and *hard work vs. natural ability* factors. However, while this same material has appeared to also help prevent prominent deterioration of *structure of science* for the modified fall students, there seems to have been essentially no notably different impact on *hard work vs. natural ability*. Given these findings, a conservative result is that the included course material has a consistent impact on student beliefs regarding the *structure of science*. Given the prevalence of views associated with this factor throughout the uncovered epistemic framework within astronomy, focusing on addressing the nature of science in astronomy holds a promising avenue for improving epistemological beliefs in general. An agressive conclusion would have us claim that we were able to impact both *structure of science* and *hard work vs. natural ability*.

There exist other interventions within physics which have been seen to improve the epistemologies of students over the course of a semester, namely the Physics and Everyday Thinking curriculum, the Modeling Instruction curriculum, and the Physics By Inquiry curriculum [24, 129, 157]. All these curricula require a complete course restructuring relying predominantly on explicit modeling (attempting to explain a question/observation using experiment guided by worksheets/tutorials, group collaboration, Socratic authorities, and limited lecture) and implicit/explicit work with the nature of science. Our work has helped define the effectiveness of explicit nature of science implementations within astronomy (particularly given the identification of a nature of science factor, i.e. *structure of science*) and precisely what epistemic constructs this type of intervention impacts. In outlining a complete epistemic framework of astronomy students, we have also contributed to how other interventions beyond those mentioned above could be directed to impact and alter specific epistemological beliefs of students within these courses.

Future Work

This research functions as a cornerstone for future work regarding the investigation of epistemological beliefs within physics and astronomy. This research is the first of its kind in the full development of an epistemic framework regarding how students are responding to one of the primary epistemological instruments in the domain of physics and astronomy. The factor *structure of science* is also the first of its kind to be observed within physics education research, further emphasizing the importance of including course material focused on the nature of science.

The volume of additional information provided by this work has led to numerous questions and avenues for future research. However, as it pertains to this study, future work includes the collection of student written responses for the EBAPS questions which were not collected due to time constraints (items 2, 13, 18, 20, and 23). Additionally, it would be ideal to collect student definitions of "understanding", "estimate", "worry", "big idea", "think more scientifically", as well as many others as it relates to their use on the EBAPS.

Other immediate future work includes the analysis of how the variables involved in this study behave as compared to each other (as opposed to how they behave across a semester). For example, direct comparisons of males to females or business students to engineering students on pretest/post-test. Whereas this study concerned itself with “within-person effects”, the work proposed would be focused on “between-persons effects.” Additionally, it would be ideal to conduct a confirmatory factor analysis on data outside of Montana State University.

Additional future work may include the modification of the EBAPS to better bring forth clear information regarding the epistemic stances taken by students responding to the instrument. As it currently exists, this researcher finds the EBAPS to do a satisfactory overall job gauging epistemological beliefs but could achieve more clarity along the axes of the instrument. Other future work could involve the development of an epistemological assessment designed to measure student epistemic views specifically within introductory astronomy. Additionally, it could be beneficial to utilize the EBAPS in an introductory physics course to develop a theoretical epistemic framework unique to specifically that context. Lastly, the implementation of nature of science material within astronomy was conservative in this study. It may be more telling if the course were to undergo a more complete immersion in material pertaining to the nature of science.

REFERENCES CITED

[1] http://macmillanlearning.com/Catalog/elearningbrowsebymediatype/SaplingLearning.

[2] AAS mission and vision statement. https://aas.org/governance/aas-mission-and-vision-statement. Accessed: 2018-09-21.

[3] Parallel analysis program for raw data and data permutations. https://people.ok.ubc.ca/brioconn/nfactors/rawpar.sps. Accessed: 2017-12-28.

[4] The r project for statistical computing. https://www.r-project.org/. Accessed: 2018-08-27.

[5] Statnotes: Topics in multivariate analysis: Factor analysis. https://www.rstudio.com/. Accessed: 2018-08-27.

[6] The idea behind ebaps. https://science.nasa.gov/learners. Accessed: 2018-09-21.

[7] Vii. mathematical contributions to the theory of evolution.—iii. regression, heredity, and panmixia. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 187:253–318, 1896.

[8] T. A. Brown. *Confirmatory Factor Analysis for Applied Research*, volume 156. 01 2006.

[9] B. A King and S. Magun-jackson. Epistemological beliefs of engineering students. 35:1–9, 01 2014.

[10] W. K. Adams, K. K. Perkins, N. S. Podolefsky, M. Dubson, N. D. Finkelstein, and C. E. Wieman. New instrument for measuring student beliefs about physics and learning physics: The colorado learning attitudes about science survey. *Phys. Rev. ST Phys. Educ. Res.*, 2:010101, Jan 2006.

[11] T. W. Adorno, E. Frenkel-Brunswik, D. Levinson, and N. Sanford. *The authoritarian personality.* Harper & Brothers, 1950.

[12] L. Aguilar, G. Walton, and C. Wieman. Psychological insights for improved physics teaching. *Physics Today*, 67:43, 2014.

[13] A. C. Aitken. Iv.on least squares and linear combination of observations. *Proceedings of the Royal Society of Edinburgh*, 55:4248, 1934.

[14] H. Akaike. Factor analysis and the aic. 52:317–332, 02 1987.

[15] D. Allchin. *Teaching the Nature of Science*. SHiPS Education Press, Saint Paul, MN, 2013.

[16] B. J. Alters. Whose nature of science? *Journal of Research in Science Teaching*, 34(1):39–55.

[17] T. A. Angelo and K. P. Cross. *Classroom assessment techniques: a handbook for college teachers*. Jossey-Bass, 1993.

[18] R. B. Kline. *Principles And Practice Of Structural Equation Modeling*. 01 2010.

[19] S. P. Bates, R. K. Galloway, C. Loptson, and K. A. Slaughter. How attitudes and beliefs about physics change from high school to faculty. *Phys. Rev. ST Phys. Educ. Res.*, 7:020114, Nov 2011.

[20] M. R. Belenky, B. M. Clinchy, N. R. Goldberger, and J. M. Tarule. *Women's ways of knowing: The development of self voice and mind*. Basic Books, 1986.

[21] I. Bernstein and G. Teng. Factoring items and factoring scales are different: Spurious evidence for multidimensionality due to item categorization. *Psychological Bulletin*, 105:467–477, 05 1989.

[22] D. Berreby. Unabsolute truths: Clifford geertz. The New York Times Magazine, 1995.

[23] T. R. Bidell and K. W. Fischer. Beyond the stage debate: Action, structure, and variability in piagetian theory and research. In R. J. Sternberg and C. A. Berg, editors, *Intellectual development*, pages 100–140. New York: Cambridge University Press, 1992.

[24] E. Brewe, L. Kramer, and G. O'Brien. Modeling instruction: Positive attitudinal shifts in introductory physics measured with class. *Phys. Rev. ST Phys. Educ. Res.*, 5(1):013102, June 2009.

[25] B. M. Byrne. *Structural Equation Modeling with AMOS: Basic Concepts, Applications, and Programming (1st Ed.)*, volume 20. 01 2001.

[26] R. C. MacCallum and S. Hong. Power analysis in covariance structure modeling using gfi and agfi. 32:193–210, 04 1997.

[27] R. C. MacCallum, M. W. Browne, and H. M. Sugawara. Power analysis and determination of sample size for covariance structure modeling. 1:130–149, 06 1996.

[28] S. G. Carmer and M. R. Swanson. An evaluation of ten pairwise multiple comparison procedures by monte carlo methods. *Journal of the American Statistical Association*, 68(341):66–74, 1973.

[29] E. G. Carmines and J. P. McIver. *Analyzing Models with Unobserved Variables: Analysis of Covariance Structures*. Sage, Beverly Hills, 1981.

[30] N. J. Castellan. Shuffling arrays: Appearances may be deceiving. *Behavior Research Methods, Instruments, & Computers*, 24:72–77, 1992.

[31] R. B. Cattell. The scree test for the number of factors. *Multivariate Behavioral Research*, 1(2):245–276, 1966.

[32] R. B. Cattell. *The scientific use of factor analysis in behavioral and life sciences.* New York ; London : Plenum Press, 1978.

[33] B. A. Cerny and H. F. Kaiser. A study of a measure of sampling adequacy for factor-analytic correlation matrices. *Multivariate Behavioral Research*, 12(1):43–47, 1977.

[34] K. Charmaz. *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis.* Sage Publications, 2006.

[35] D. Child. The essentials of factor analysis. 3, 01 2006.

[36] M.-H. Cho, D. Lankford, and D. Wescott. Exploring the relationships among epistemological beliefs, nature of science, and conceptual change in the learning of evolutionary theory. *Evolution Education and Outreach*, 4:313–322, 06 2011.

[37] C. P. Chou and P. M. Bentler. Estimates and tests in structural equation modeling. In *Structural equation modeling: Issues and application*, pages 37–55. Newbury, CA: SAGE publications, 1995.

[38] C.-P. Chou, P. M Bentler, and A. Satorra. Scaled test statistics and robust standard errors for non-normal data in covariance structure analysis: A monte carlo study. 44 (Pt 2):347–57, 12 1991.

[39] R. Coe. It's the effect size, stupid: What effect size is and why it is important, 09 2002.

[40] J. Cohen. Statistical power analysis for the behavioral sciences. 2nd, 01 1988.

[41] J. Cohen and J. Cohen. *Applied multiple regression/correlation analysis for the behavioral sciences.* L. Erlbaum Associates, 2003.

[42] A. Collins and W. Ferguson. Epistemic forms and epistemic games: Structures and strategies to guide inquiry. *Educational Psychologist*, 28:25–42, 1993.

[43] A. L. Comrey and H. B. Lee. *A first course in factor analysis.* Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc., 2 edition, 1992.

[44] W. J. Conover. *Practical nonparametric statistics.* Academic Internet Publishers, 2007.

[45] J. Corbin and A. Strauss. *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory.* Thousand Oaks, CA: Sage, 3 edition, 2008.

[46] D. Cosnette. The apollo hoax.

[47] A. B. Costello and J. Osborne. Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*, 10, 2005.

[48] N. R. Council. *National Science Education Standards.* The National Academies Press, Washington, DC, 1996.

[49] M. G. R. Courtney. Determining the number of factors to retain in efa : Using the spss r-menu v 2 . 0 to make more judicious estimations. 2015.

[50] G. Cumming and S. Finch. Inference by eye: Confidence intervals and how to read pictures of data. *American Psychologist*, pages 170–180, 2005.

[51] J. DeCoster. Overview of factor analysis. 01 1998.

[52] G. Deming and B. Hufnagel. Who's taking astro 101? *The Physics Teacher*, 39(6):368–369, 2001.

[53] J. Dewey. *How we think: A restatement of the relation of reflective thinking to the educative process.* Lexington, MA: Heath, 1933.

[54] L. Ding and K. Mollohan. How college-level introductory instruction can impact student epistemological beliefs. *Journal of College Science Teaching*, 44, 2015.

[55] L. Ding and P. Zhang. Making of epistemologically sophisticated physics teachers: A cross-sequential study of epistemological progression from preservice to in-service teachers. *Phys. Rev. Phys. Educ. Res.*, 12:020137, Nov 2016.

[56] L. C. Dinneen and B. C. Blakesley. [algorithm as 62] a generator for the sampling distribution of the mann-whitney u statistic. *Applied Statistics*, 22:269–273, 1973.

[57] A. A. diSessa. Towards an epistemology of physics. *Cognition and Instruction*, 10:105–225, 1993.

[58] K. A. Douglas, M. S. Yale, D. E. Bennett, M. P. Haugan, and L. A. Bryan. Evaluation of colorado learning attitudes about science survey. *Phys. Rev. ST Phys. Educ. Res.*, 10:020128, Nov 2014.

[59] G. Duffy, S. Chance, and B. Bowe. Improving engineering students design skills in a project-based learning course by addressing epistemological issues. In *Proceedings of the 40th SEFI Conference*. AIP Conference Proceedings, 2012.

[60] D. Duncan and L. Arthurs. Improving student attitudes about learning science and student scientific reasoning skills. *Astronomy Education Review*, 11:010102, 2012.

[61] D. B. Duncan. Multiple range and multiple f tests. *Biometrics*, 11(1):1–42, 1955.

[62] C. S. Dweck. *Mindset: The new psychology of success*. New York: Random House, 2006.

[63] C. S. Dweck and J. Bempechat. Children's theories of intelligence: Consequences for learning. In S. G. Paris, G. M. Olson, and H. W. Stevenson, editors, *Learning and motivation in the classroom*, pages 239–256. Hillsdale, NJ: Earlbaum, 1983.

[64] C. S. Dweck and E. L. Leggett. A social-cognitive approach to motivation and personality. *Psychological Review*, 95:256–273, 1988.

[65] A. E. Hendrickson and P. Owen White. Promax: A quick method for rotation to oblique simple structure. *British Journal of Statistical Psychology*, 17:65–70, 05 1964.

[66] G. E. Schwarz. Estimating the dimension of a model. 6, 03 1978.

[67] I. Einot and K. R. Gabriel. A study of the powers of several methods of multiple comparisons. 70, 09 1975.

[68] A. Elby. Epistemological beliefs assessment for physical science (ebaps). `http://www2.physics.umd.edu/~elby/EBAPS/home.htm`. Accessed: 2018-05-29.

[69] A. Elby. The idea behind ebaps. `http://www2.physics.umd.edu/~elby/EBAPS/idea.htm`. Accessed: 2018-05-29.

[70] A. Elby. Another reason that physics students learn by rote. *American Journal of Physics*, 67(1):52–57, 1999.

[71] A. Elby. Getting started with research on epistemologies and expectations. In *Getting Started in PER*, volume 2. 1 edition, October 2010.

[72] E. Erikson. Identity and the life cycle: Selected papers. *Psychological Issues*, 1:1–171, 1959.

[73] H. F. Kaiser. The varimax criterion for analytic rotation in factor analysis. 23:187–200, 02 1958.

[74] H. F. Kaiser. An index of factorial simplicity. *Psychometrika*, 39:31–36, 02 1974.

[75] L. R. Fabrigar, D. T. Wegener, R. C. MacCallum, and E. J. Strahan. Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4:272–299, 1999.

[76] S. Finney and C. DiStefano. Non-normal and categorical data in structural equation modeling. pages 269–314, 01 2013.

[77] K. W. Fischer and S. L. Pipp. Processes of cognitive development: Optimal level and skill acquisition. In R. J. Sternberg, editor, *Mechanisms of cognitive development*, pages 45–80. New York: Freeman, 1984.

[78] R. A. Fisher. Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika*, 10(4):507–521, 1915.

[79] R. A. Fisher. On the probable error of a coefficient of correlation deduced from a small sample. *Metron*, 1:3–32, 1921.

[80] R. A. Fisher. The distribution of the partial correlation coefficient. *Metron*, 3:329–332, 1924.

[81] R. A. Fisher. *Statistical Methods for Research Workers*, pages 66–70. Springer New York, New York, NY, 1992.

[82] P. W. Foos. Effects of student-written questions on student test performance. *Teaching of Psychology*, 16(2):77–78, 1989.

[83] J. R. Fraenkel and N. E. Wallen. *How to Design and Evaluate Research in Education*. McGraw-Hill, 1996.

[84] A. Fraknoi. Enrollments in astronomy 101 courses. 1, 10 2001.

[85] S. Freeman, S. L. Eddy, M. McDonough, M. K. Smith, N. Okoroafor, H. Jordt, and M. P. Wenderoth. Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, 111(23):8410–8415, 2014.

[86] S. G. West, J. F. Finch, and P. J. Curran. Structural equation models with non-normal variables: Problems and remedies. 01 1995.

[87] R. K. Gable and M. B. Wolf. *Instrument Development in the Affective Domain*. Springer Netherlands, 2 edition, 1993.

[88] D. Garson. Statnotes: Topics in multivariate analysis: Factor analysis. `https://faculty.chass.ncsu.edu/garson/PA765/statnote.htm`. Accessed: 2017-12-29.

[89] C. Gilligan. *In a different voice: Psychological theory and women's development.* Harvard University Press, 1982.

[90] B. Glaser and A. Strauss. *The discovery of grounded theory: Strategies for qualitative research.* Aldine, 1967.

[91] E. Guadagnoli and W. Velicer. Relation of sample size to the stability of component patterns. 103:265–75, 04 1988.

[92] K. Hakan and E. Munire. Academic motivation: Gender, domain and grade differences. 143, 08 2014.

[93] D. Hammer. Two approaches to learning physics. *Phys. Teach.*, 27(9):664–670, 1989.

[94] D. M. Hammer. Epistemological beliefs in introductory physics. *Cognition and Instruction*, 12:151–183, 1994.

[95] D. M. Hammer. Student resources for learning introductory physics. *American Journal of Physics*, 68:52–59, 2000.

[96] D. M. Hammer and A. Elby. Tapping epistemological resources for learning physics. *Journal of the Learning Sciences*, 12:53–90, 2003.

[97] D. M. Hammer, A. Elby, A. Scherr, and E. F. Redish. Resources, framing, and transfer. In *Transfer of learning from a modern multidisciplinary perspective*, pages 89–120. Greenwich, CT: Information Age Publishing, 2005.

[98] H. H. Harman. *Modern Factor Analysis.* University of Chicago press, 3rd ed. edition, 1976.

[99] C. Harris. Note on the squared multiple correlation as a lower bound to communality. *Psychometrika*, 43(2):283–284, June 1978.

[100] Z. Hazari, R. H. Tai, and P. M. Sadler. Gender differences in introductory university physics performance: The influence of high school physics preparation and affective factors. *Science Education*, 91(6):847–876, 2007.

[101] A. Heikkil and K. Lonka. Studying in higher education: students' approaches to learning, selfregulation, and cognitive strategies. *Studies in Higher Education*, 31(1):99–117, 2006.

[102] M. Hennessey. Metacognitive aspects of students' reflective discourse: Implications for intentional conceptual change teaching and learning. pages 103–132, 01 2003.

[103] P. Hewson, M. Beeth, and N. Thorley. Teaching for conceptual change. pages 199–218, 1998.

[104] B. K. Hofer and P. R. Pintrich. The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, 67(1):88–140, 1997.

[105] T. Hong, E. Purzer, and M. E. Cardella. A psychometric re-evaluation of the design, engineering and technology (det) survey. *Journal of Engineering Education*, 100(4):800–818, 2011.

[106] D. Hooper, J. Coughlan, and M. R. Mullen. Structural equation modeling: Guidelines for determining model fit. 6, 11 2007.

[107] H. Hotelling. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24:417–441, 1933.

[108] H. Hotelling. Relations between two sets of variates. *Biometrika*, 28:321–377, 1936.

[109] M. C. Howard. A review of exploratory factor analysis decisions and overview of current practices: What we are doing and how can we improve? *International Journal of HumanComputer Interaction*, 32(1):51–62, 2016.

[110] L. Hu, P. M. Bentler, and Y. Kano. Can test statistics in covariance structure analysis be trusted? 112:351–62, 10 1992.

[111] L.-t. Hu and P. M Bentler. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. 6:1–55, 01 1999.

[112] D. J. Bartholomew, M. Knott, and I. Moustaki. *Latent Variable Models and Factor Analysis: A Unified Approach*. 06 2011.

[113] P. J. Curran, S. G. West, and J. F. Finch. The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. 1:16–29, 03 1996.

[114] R. I. Jennrich and P. F. Sampson. Rotation for simple loadings. 31:313–323, 09 1966.

[115] K. Johnson and S. Willoughby. Changing epistemological beliefs with nature of science implementations. *Phys. Rev. Phys. Educ. Res.*, 14:010110, Feb 2018.

[116] K. Johnson and S. D. Willoughby. Epistemic belief structures within introductory astronomy. *Phys. Rev. Phys. Educ. Res.*, 14:010135, Jun 2018.

[117] K. Jreskog and D. Srbom. *LISREL-VI user's guide (3rd ed.)*. Mooresville, IN: Scientific Software, 1984.

[118] H. F. Kaiser. The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, 20:141–151, Apr 1960.

[119] D. Kaplan. *Structural Equation Modeling. Foundations and Extensions*, volume 10. 01 2000.

[120] P. M. King and K. Strohm-Kitchener. *Developing reflective judgment: Understanding and promoting intellectual growth and critical thinking in adolescents and adults*. San Francisco: Jossey-Bass, 1994.

[121] K. Kitchener. Cognition, metacognition, and epistemic cognition: A three-level model of cognitive processing. 26:222–232, 01 1983.

[122] L. Kohlberg. Stage and sequence: The cognitive-developmental approach to socialization. In D. A. Goslin, editor, *Handbook of socialization theory and research*, pages 347–480. New York: Rand McNally, 1969.

[123] W. Krijnen. Algorithms for unweighted least-squares factor analysis. 21:133–147, 02 1996.

[124] T. Kubiszyn and G. Borich. *Educational Testing and Measurement*. Wiley, New York, 2007.

[125] D. Kuhn. *The Skill of Argument*. Cambridge University Press, 1991.

[126] M. D. LeCompte and J. Preissle. *Ethnography and Qualitative Design in Educational Research*. San Diego: Academic Press, 2 edition, 1993.

[127] N. Lederman. Students' and teachers' conceptions of the nature of science: A review of the research. 29:331–359, 04 1992.

[128] N. Lederman, F. Abd-El-Khalick, R. L. Bell, and R. S. Schwartz. Views of nature of science questionnaire: Toward valid and meaningful assessment of learners conceptions of nature of science. *J. Res. Sci. Teaching*, 39(6):497–521, July 2002.

[129] B. Lindsey, L. Hsu, H. Sadaghiani, J. Taylor, and K. Cummings. Positive attitudinal shifts with the physics by inquiry curriculum across multiple implementations. *Phys. Rev. ST Phys. Educ. Res.*, 8(1):010102, January 2012.

[130] L. Lising and A. Elby. The impact of epistemology on learning: A case study from introductory physics. *American Journal of Physics*, 73:372–382, 2005.

[131] L. Louca, A. Elby, D. M. Hammer, and T. Kagey. Epistemological resources: Applying a new epistemological framework to science instruction. *Educational Psychologist*, 39:57–68, 2004.

[132] P. M. Bentler. Comparative fit indexes in structural models. 107:238–46, 04 1990.

[133] P. M. Bentler and D. Bonett. Significance tests and goodness-of-fit in analysis of covariance structures. 88:588–606, 11 1980.

[134] P. M. Bentler and P. Dudgeon. Covariance structure analysis: Statistical practice, theory, and directions. 47:563–92, 02 1996.

[135] M. B. Magolda. *Knowing and Reasoning in College: Gender-Related Patterns in Students' Intellectual Development*. San Francisco: Jossey-Bass, 1 edition, 1992.

[136] A. Maltese and R. Tai. Eyeballs in the fridge: Sources of early interest in science. 32:669–685, 03 2010.

[137] H. B. Mann and D. R. Whitney. On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*, 18(1):50–60, 1947.

[138] H. Marsh and D. Hocevar. Application of confirmatory factor analysis to the study of self-concept. first- and higher order factor models and their invariance across groups. 97:562–582, 05 1985.

[139] J. Marx. Attitudes of undergraduate general science students toward learning science and the nature of science. In *AIP Conference Proceedings*. 2005.

[140] J. Marx and K. Cummings. Normalized change. *Am. J. Phys.*, 75(1):87–91, 2007.

[141] R. E. Mayer and S. A. Ambrose. *How learning works: seven research-based principles for smart teaching*. Jossey-Bass, 2010.

[142] W. F. McComas and J. K. Olson. The nature of science in international science education standards documents. In *The Nature of Science in Science Education: Rationales and Strategies*, pages 41–52. Kluwer Academic Publishers, 1998.

[143] R. P. McDonald. A simple comprehensive model for the analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology*, 31(1):59–72, 1978.

[144] C. Mcintosh. Rethinking fit assessment in structural equation modelling: A commentary and elaboration on barrett (2007). 42:859–867, 05 2007.

[145] W. McKeachie and M. Svinicki. How to make lectures more effective. In *McKeachie's Teaching Tips*, pages 58–73. 2006.

[146] V. Mellado. Preservice teachers' classroom practice and their conceptions of the nature of science. 6:331–354, 07 1997.

[147] M. B. Miles and A. M. Huberman. *Qualitative data analysis: an expanded sourcebook*. Thousand Oaks, CA: Sage, 1994.

[148] J. D. Miller. The measurement of civic scientific literacy. *Public Understanding of Science*, 7(3):203–223, 1998.

[149] D. Muchoney. *The implementation of an interactive engagement model of instruction in high school physics*. PhD thesis, The Robert Morris University, 2008.

[150] B. Muthén. A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika*, 49(1):115–132, Mar 1984.

[151] J. Neuhaus and C. Wrigley. The quartimax method: An analytic approach to orthogonal simple structure. 7:81–91, 1954.

[152] J. D. Novak. *Learning, creating, and using knowledge: concept maps as facilitative tools in schools and corporations*. Routledge, 2012.

[153] B. P. O'Connor. Spss and sas programs for determining the number of components using parallel analysis and velicers map test. *Behavior Research Methods, Instruments, & Computers*, 32:396–402, 2000.

[154] U. of Maryland PERG. Student expectations in university physics. `https://www.physics.umd.edu/perg/expects/index.html`, 2018. Accessed: 2018-04-04.

[155] U. Olsson. Maximum likelihood estimation of the polychoric correlation coefficient. *Psychometrika*, 44(4):443–460, Dec 1979.

[156] J. Osborne and D. C. Fitzpatrick. Replication analysis in exploratory factor analysis: What it is and why it makes your analysis better. *Practical Assessment, Research & Evaluation*, 17, 2012.

[157] V. K. Otero and K. E. Gray. Learning to think like scientists with the pet curriculum. volume 951, pages 160–163. AIP Conference Proceedings, 2007.

[158] R. P. Mcdonald. *Factor Analysis and Related Methods*. Lawrence Erlbaum Associates, Inc, 1985.

[159] B. Partridge and G. Greenstein. Goals for "astro 101": Report on workshops for department leaders. *Astronomy Education Review*, 2:46–89, 2004.

[160] E. Pascarella and P. Terenzini. *How college affects students: A third decade of research*. San Francisco: Jossey-Bass, 1991.

[161] M. B. Paulsen and C. T. Wells. Domain differences in the epistemological beliefs of college students. *Research in Higher Education*, 39(4):365–384, 1998.

[162] K. Pearson. X. on the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 50(302):157–175, 1900.

[163] K. Pearson. Liii. on lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2(11):559–572, 1901.

[164] J. W. Pellegrino. Assessment of science learning: Living in interesting times. *Journal of Research in Science Teaching*, 49(6):831–841, 2012.

[165] W. G. Perry. *Forms of Ethical and Intellectual Development in the College Years: A Scheme*. Rinehart and Winston, 1970.

[166] PhysPort. Beliefs/attitudes assessments. `https://www.physport.org/assessments/?Focus=140`, 2018. Accessed: 2018-04-03.

[167] J. Piaget. Intellectual evolution from adolescence to adulthood. *Human Development*, 15:1–12, 1972.

[168] E. E. Prather, T. F. Slater, J. P. Adams, and G. Brissenden. *Lecture-Tutorials for Introductory Astronomy*. Pearson, 2012.

[169] P. Proitsi, G. Hunter, M. Tsolaki, M. Lupton, M. Daniilidou, P. Hollingworth, N. Archer, C. Foy, F. Stylios, B. Mcguinness, S. Todd, B. Lawlor, M. Gill, C. Brayne, D. Rubinsztein, M. Owen, J. Williams, D. Craig, P. Passmore, and J. Powell. A multiple indicators multiple causes (mimic) model of behavioural and psychological symptoms in dementia (bpsd). 32:434–42, 05 2009.

[170] P. R. Pintrich, C. A. Wolters, and G. Baxter. Assessing metacognition and selfregulated learning. pages 43–98, 01 2000.

[171] E. F. Redish. A theoretical framework for physics education research: Modeling student thinking. In *Proceedings of the Enrico Fermi Summer School, Course CLVI*, pages 1–63. Bologna: Italian Physical Society, 2004.

[172] E. F. Redish, J. M. Saul, and R. N. Steinberg. Student expectations in introductory physics. *American Journal of Physics*, 66(3):212–224, 1998.

[173] M. Rice and G. Harris. Comparing effect sizes in follow-up studies: Roc area, cohen's d, and r. 29:615–20, 11 2005.

[174] K. F. Riley, M. P. Hobson, and S. J. Bence. *Mathematical methods for physics and engineering*. Cambridge University Press, 3rd ed edition, 2006.

[175] R. J. Rummel. *Applied Factor Analysis*. Northwestern University Press, 1970.

[176] D. W. Russell. In search of underlying dimensions: The use (and abuse) of factor analysis in personality and social psychology bulletin. 28:1629–1646, 2002.

[177] M. P. Ryan. Monitoring text comprehension: Individual differences in epistemological standards. *Journal of Educational Psychology*, 76:249–258, 1984.

[178] M. Sackes. *The Role of Cognitive, Metacognitive, and Motivational Variables in Conceptual Change: Preservice Early Childhood Teachers' Conceptual Understanding of the Cause of Lunar Phases*. PhD thesis, 2010.

[179] C. Sagan. Why we need to understand science. *Mercury*, 22(2):52, Apr 1993.

[180] A. Schoenfeld. When good teaching leads to bad results: The disasters of "well taught" mathematics classes. *Educational Psychologist*, 23:145–166, 1988.

[181] A. H. Schoenfeld. Beyond the purely cognitive: Belief systems, social cognitions, and metacognitions as driving forces in intellectual performance. *Cognitive Science*, 7:329–363, 1983.

[182] A. H. Schoenfeld. *Mathematical problem solving*. San Diego, CA: Academic Press, 1985.

[183] M. Schommer. Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology*, 85:498–504, 1990.

[184] M. Schommer. Epistemological development and academic performance among secondary students. *Journal for Educational Psychology*, 85:406–411, 1993.

[185] M. Schommer, A. Crouse, and N. Rhodes. Epistemological beliefs and mathematical text comprehension: Believing it's simple doesn't make it so. *Journal for Educational Psychology*, 84:435–443, 1992.

[186] M. Schommer and P. A. Dunnell. A comparison of epistmological beliefs between gifted and nongifted high school students. *Roeper Review*, 16:207–210, 1994.

[187] G. Schraw, M. E. Dunkle, and L. D. Bendixen. Cognitive processes in well-defined and ill-defined problem solving. *Applied Cognitive Psychology*, 9:523–538, 1995.

[188] S. Sharma, S. Mukherjee, A. Kumar, and W. Dillon. A simulation study to investigate the use of cutoff values for assessing model fit in covariance structure models. 58:935–943, 02 2005.

[189] M. Shevlin and J. Miles. Effects of sample size, model specification and factor loadings on the gfi in confirmatory factor analysis. 25:85–90, 07 1998.

[190] T. Slater, J. P. Adams, G. Brissenden, and D. Duncan. What topics are taught in introductory astronomy courses? *The Physics Teacher*, 39(1):52–55, 2001.

[191] C. Spearman. "general intelligence," objectively determined and measured. *The American Journal of Psychology*, 15(2):201–292, 1904.

[192] G. G. Stern. *Inventory of beliefs*. Chicago: University of Chicago Press, 1953.

[193] A. Strauss and J. Corbin. Grounded theory methodology: An overview. In N. K. Denzin and Y. S. Lincoln, editors, *Handbook of qualitative research*, pages 273–285. Thousand Oaks, CA: Sage, 1994.

[194] M. D. Sundberg, M. L. Dini, and E. Li. Decreasing course content improves student comprehension of science and attitudes towards science in freshman biology. *Journal of Research in Science Teaching*, 31(6):679–693, 1994.

[195] J. Sweller. Cognitive load during problem solving: Effects on learning. 12:257–285, 04 1988.

[196] B. G. Tabachnick and L. S. Fidell. *Using Multivariate Statistics*. Allyn & Bacon, Inc., Needham Heights, MA, USA, 2001.

[197] S. J. Taylor and R. Bogdan. *Introduction to Qualitative Research Methods*. Wiley, 3 edition, 1998.

[198] R. Tesch. *Qualitative research: Analysis types and software tools*. Psychology Press, 1990.

[199] L. L. Thurstone. *Multiple-factor analysis: a development and expansion of the vectors of mind*. University of Chicago press, 1947.

[200] M. E. Timmerman and U. Lorenzo-Seva. Dimensionality assessment of ordered polytomous items with parallel analysis. *Psychological Methods*, 16:209–220, 2011.

[201] K. Tobin, D. J. Tippins, and K. Hook. Referents for changing a science curriculum: A case study of one teacher's change in beliefs. 3:245–264, 07 1994.

[202] L. R. Tucker and R. C. MacCullum. *Exploratory Factor Analysis*. 1997.

[203] J. W. Tukey. Comparing individual means in the analysis of variance. *Biometrics*, 5(2):99–114, 1949.

[204] H. Vilppu, M. Mikkil-Erdmann, and I. Ahopelto. The role of regulation and processing strategies in understanding science text among university students. *Scandinavian Journal of Educational Research*, 57(3):246–262, 2013.

[205] L. S. Vygotsky. *Thought and language*. Cambridge, MA: MIT Press, 1962.

[206] A. Wald. Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Transactions of the American Mathematical Society*, 54(3):426–482, 1943.

[207] R. A. Waller and D. B. Duncan. A bayes rule for the symmetric multiple comparisons problems. *Journal of the American Statistical Association*, 64(328):1484–1503, 1969.

[208] A. J. Wayne and P. Youngs. Teacher characteristics and student achievement gains: A review. *Review of Educational Research*, 73(1):89–122, 2003.

[209] B. Wheaton, B. Muthn, D. F. Alwin, and G. F. Summers. Assessing reliability and stability in panel models. *Sociological Methodology*, 8:84–136, 1977.

[210] F. Wilcoxon. Individual comparisons by ranking methods. *Biometrics Bulletin*, 1(6):80–83, 1945.

[211] S. S. Wilks. The large-sample distribution of the likelihood ratio for testing composite hypotheses. *Ann. Math. Statist.*, 9(1):60–62, 03 1938.

[212] S. D. Willoughby and K. Johnson. Epistemic beliefs of non-stem majors regarding the nature of science: Where they are and what we can do. *American Journal of Physics*, 85(6):461–468, 2017.

[213] H. F. Wolcott. *Transforming qualitative data: description, analysis, and interpretation*. Thousand Oaks, CA: Sage, 1994.

[214] H. Yanai and M. Ichikawa. New lower and upper bounds for communality in factor analysis. *Psychometrika*, 55(2):405409, 1990.

APPENDICES

APPENDIX A

EBAPS AXES AND EFA FACTOR LOADINGS

EBAPS Axis Summary

Table A.1: Description of EBAPS Axes and the Items Associated to These Axes. These are the definitions as provided by the EBAPS homepage [68]

1.	Structure of scientific knowledge (Items: 2, 8, 10, 15, 17, 19, 20, 23, 24, 28) Is [science] knowledge a bunch of weakly connected pieces without much structure and consisting mainly of facts and formulas? Or is it a coherent, conceptual, highly-structured, unified whole?
2.	Nature of knowing and learning (Items: 1, 7, 11, 12, 13, 18, 26, 30) Does learning science consist mainly of absorbing information? Or, does it rely crucially on constructing one's own understanding by working through the material actively, by relating new material to prior experiences, intuitions, and knowledge, and by reflecting upon and monitoring one's understanding?
3.	Real-life applicability (Items: 3, 14, 19, 27) Are scientific knowledge and scientific ways of thinking applicable only to restricted spheres such as the classroom or the laboratory? Or, does science apply more generally to real life? These items tease out students' views of the applicability of scientific knowledge as distinct from the student's own desire to apply science to real life, which depends on the student's interests, goals, and other non-epistemological factors.
4.	Evolving knowledge (Items: 6, 28, 29) This dimension probes the extent to which students navigate between the twin perils of absolutism (thinking all scientific knowledge is set in stone) and extreme relativism (making no distinctions between evidence-based reasoning and mere opinion).
5.	Source of ability to learn (Items: 5, 9, 16, 22, 25) Is being good at science mostly a matter of fixed natural ability? Or, can most people become better at learning (and doing) science? As much as possible, these items probe students' epistemological views about the efficacy of hard work and good study strategies, as distinct from their self-confidence and other beliefs about themselves.

Additional Factor Analysis Loading Tables

Table A.2: Finalized Factor Loadings for Group A. Loadings less than 0.25 have been suppressed. EFA involved approximately half of the section 1 students. These loadings represent results after a filtering of the original 30 items and a comparison with group B.

Items	Factor 1	Factor 2	Factor 3
05	0.645		
16	0.576		
09	0.485		
11	0.431		
12	0.419		
22		0.733	
21		0.618	
25		0.403	
08			-0.644
14			-0.598
06			-0.496
10			-0.481

Table A.3: Finalized Factor Loadings for Group B. Loadings less than 0.25 have been suppressed. EFA involved approximately half of the section 1 students. These loadings represent results after a filtering of the original 30 items and a comparison with group A.

Items	Factor 1	Factor 2	Factor 3
16	0.596		
11	0.498		
05	0.491		
12	0.460		
09	0.438		
22		0.660	
21		0.637	
25		0.505	
08			0.572
14			0.556
10			0.474
06			0.467

Table A.4: Finalized 5-Factor Loadings for Group A. Loadings less than 0.25 have been suppressed. EFA involved approximately half of the section 1 students. These loadings represent results after a filtering of the original 30 items and a comparison with group B.

Items	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
08	0.586				
14	0.579				
06	0.505				
10	0.436				
22		0.781			
21		0.609			
25		0.328			
30			0.520		
28			0.444		
11				0.551	
05				0.525	
16				0.413	
12				0.363	
09					-0.513
04	0.304				-0.482

Table A.5: Finalized 5-Factor Loadings for Group B. Loadings less than 0.25 have been suppressed. EFA involved approximately half of the section 1 students. These loadings represent results after a filtering of the original 30 items and a comparison with group A.

Items	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
16	0.582				
05	0.541				
09	0.498				
11	0.427				
12	0.402				
22		0.715			
21		0.585			
25		0.471			
04			0.654		
30				0.600	
28				0.378	
06					0.492
14					0.480
08					0.437
10					0.412

APPENDIX B

DATA DESCRIPTIVES FOR EFA IN R

Descriptive Tables

Table B.1: Descriptive Statistics for all EBAPS Section 1 Post-Test Data. 1233 student responses in total.

Question	n	mean	s.d.	median	skew	kurtosis	s.e.
Q01	1232	2.46	1.37	3.0	-0.45	-1.36	0.04
Q02	1233	1.13	1.23	1.5	0.75	-0.46	0.03
Q03	1231	2.85	1.23	3.5	-0.99	-0.30	0.04
Q04	1231	2.57	1.12	3.0	-0.38	-0.74	0.03
Q05	1231	3.13	1.11	3.0	-1.46	1.46	0.03
Q06	1231	3.13	1.24	4.0	-0.95	-0.55	0.04
Q07	1233	2.60	1.23	3.0	-0.51	-0.80	0.03
Q08	1233	2.96	1.15	3.0	-1.08	0.17	0.03
Q09	1233	2.81	1.12	3.0	-0.86	-0.04	0.03
Q10	1231	2.65	1.15	3.0	-0.52	-0.68	0.03
Q11	1232	2.52	1.11	3.0	-0.68	-0.22	0.03
Q12	1233	2.64	1.27	3.0	-0.70	-0.89	0.04
Q13	1231	1.32	1.22	1.0	0.95	-0.57	0.03
Q14	1233	3.03	1.12	3.0	-1.08	0.35	0.03
Q15	1232	2.22	1.02	2.0	-0.08	-0.56	0.03
Q16	1233	2.90	1.09	3.0	-1.01	0.40	0.03
Q17	1232	2.02	1.22	1.5	0.07	-1.27	0.03
Q18	1232	2.55	1.29	3.5	-0.45	-1.36	0.04
Q19	1232	2.29	1.55	3.0	-0.39	-1.39	0.04
Q20	1229	2.53	1.33	3.0	-0.59	-0.81	0.04
Q21	1232	3.20	0.88	3.0	-0.84	-0.03	0.03
Q22	1232	3.02	1.02	3.0	-0.81	-0.07	0.03
Q23	1231	1.69	1.27	1.0	0.46	-0.84	0.04
Q24	1230	2.97	1.38	4.0	-0.77	-1.05	0.04
Q25	1231	2.71	1.47	4.0	-0.48	-1.37	0.04
Q26	1228	3.35	1.16	4.0	-1.46	0.68	0.03
Q27	1228	3.32	1.14	4.0	-1.26	0.12	0.03
Q28	1228	2.79	1.19	3.0	-0.69	-0.54	0.03
Q29	1220	1.83	1.56	2.0	0.15	-1.35	0.04
Q30	1218	2.75	1.33	3.0	-0.70	-0.76	0.04

Table B.2: Measures of Sampling Adequacy for the Correlation Matrix of Section 1 EBAPS Data. Overall MSA = 0.850.

Q01	Q02	Q03	Q04	Q05	Q06	Q07	Q08	Q09	Q10
0.818	0.853	0.893	0.869	0.865	0.904	0.876	0.890	0.888	0.919
Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20
0.828	0.869	0.610	0.908	0.823	0.894	0.833	0.804	0.742	0.736
Q21	Q22	Q23	Q24	Q25	Q26	Q27	Q28	Q29	Q30
0.719	0.741	0.568	0.809	0.841	0.768	0.809	0.849	0.642	0.860

APPENDIX C

CFA ESTIMATES

Informational Tables for Confirmatory Factor Analysis

Table C.1: Unstandardized and Standardized Factor Loading Estimates for the 4-Factor Model with ML Estimator
*** $p < .001$

Question	Factor	Unstd. Est.	Std. Est.
Q14	Structure of Science	0.688***	0.634
Q10	Structure of Science	0.380***	0.334
Q08	Structure of Science	0.755***	0.642
Q06	Structure of Science	0.521***	0.409
Q25	Hard Work vs. Natural Ability	0.805***	0.566
Q22	Hard Work vs. Natural Ability	0.558***	0.557
Q21	Hard Work vs. Natural Ability	0.386***	0.428
Q16	Source of Ability to Learn	0.731***	0.665
Q12	Source of Ability to Learn	0.662***	0.496
Q09	Source of Ability to Learn	0.730***	0.644
Q05	Source of Ability to Learn	0.689***	0.617
Q27	Evaluating Uncertainty	0.530***	0.445
Q28	Evaluating Uncertainty	0.610***	0.510
Q30	Evaluating Uncertainty	0.593***	0.435

Table C.2: Unstandardized and Standardized Factor Loading Estimates for the 5-Factor Model with ML Estimator
*** $p < .001$

Question	Factor	Unstd. Est.	Std. Est.
Q14	Structure of Science	0.688***	0.634
Q10	Structure of Science	0.383***	0.337
Q08	Structure of Science	0.759***	0.646
Q06	Structure of Science	0.512***	0.401
Q25	Hard Work vs. Natural Ability	0.795***	0.559
Q22	Hard Work vs. Natural Ability	0.564***	0.564
Q21	Hard Work vs. Natural Ability	0.391***	0.433
Q16	Source of Ability to Learn	0.752***	0.684
Q12	Source of Ability to Learn	0.391***	0.293
Q09	Source of Ability to Learn	0.758***	0.669
Q05	Source of Ability to Learn	0.687***	0.615
Q27	Evaluating Uncertainty	0.528***	0.443
Q28	Evaluating Uncertainty	0.612***	0.512
Q30	Evaluating Uncertainty	0.593***	0.435
Q01	Science to Self	0.564***	0.417
Q12	Science to Self	0.506***	0.380
Q26	Science to Self	0.376***	0.330

Table C.3: Covariance Estimates for the 4-Factor Model with ML Estimator
*** $p < .001$
** $p < .01$
* $p < .05$

Construct	Construct	Cov. Est.
Structure of Science	Source of Ability to Learn	0.812***
Hard Work vs. Natural Ability	Source of Ability to Learn	0.355***
Source of Ability to Learn	Evaluating Uncertainty	0.461***
Structure of Science	Hard Work vs. Natural Ability	0.254***
Structure of Science	Evaluating Uncertainty	0.549***
Hard Work vs. Natural Ability	Evaluating Uncertainty	0.498***
e22	e21	0.197***
e16	e09	0.031
e10	e06	0.141***

Table C.4: Covariance Estimates for the 5-Factor Model with ML Estimator
*** $p < .001$
** $p < .01$
* $p < .05$

Construct	Construct	Cov. Est.
Structure of Science	Hard Work vs. Natural Ability	0.253***
Structure of Science	Evaluating Uncertainty	0.549***
Structure of Science	Source of Ability to Learn	0.777***
Structure of Science	Science to Self	0.603***
Hard Work vs. Natural Ability	Evaluating Uncertainty	0.497***
Hard Work vs. Natural Ability	Source of Ability to Learn	0.367***
Hard Work vs. Natural Ability	Science to Self	0.163*
Evaluating Uncertainty	Source of Ability to Learn	0.436***
Evaluating Uncertainty	Science to Self	0.417***
Source of Ability to Learn	Science to Self	0.440***
e22	e21	0.191***
e16	e09	-0.005
e10	e06	0.143***

APPENDIX D

ADDITIONAL INFORMATION FOR COURSE MODIFICATIONS STUDY

Informational Tables Pre-Post Analysis

Table D.1: Pre-test mean across all study semesters on EBAPS overall axis. Scores range from 0 to 4.

Semester	N	Pre-test	s.d.
F12	231	2.61	0.33
S13	247	2.61	0.36
F13	194	2.60	0.36
S14	234	2.61	0.37
F14	273	2.67	0.35
S15	259	2.62	0.38
F15	264	2.70	0.37
S16	134	2.62	0.39
F16	267	2.70	0.33
S17	232	2.64	0.38

Table D.2: Duncan post-hoc results on Pre-test EBAPS averages. Scores range from 0 to 4.

Semester	N	Group 1	Group 2
F13	194	2.60	
S13	247	2.60	
S14	234	2.61	
F12	231	2.62	
S16	134	2.62	
S15	259	2.62	
S17	232	2.64	2.64
F14	273	2.67	2.67
F16	267		2.70
F15	264		2.70
Sig.		0.059	0.075

Table D.3: Descriptives for Normalized Change by Population. Scores range from 0 to 4. Axis 1: Structure of scientific knowledge, Axis 2: Nature of knowing and learning, Axis 3: Real-life applicability, Axis 4: Evolving knowledge, Axis 5: Source of ability to learn

	Population	N	Mean	s.d.
Axis 1	Baseline	906	0.032	0.23
	Modified Spring	624	0.051	0.23
	Modified Fall	803	0.025	0.22
Axis 2	Baseline	906	0.018	0.27
	Modified Spring	624	0.030	0.26
	Modified Fall	803	0.013	0.25
Axis 3	Baseline	890	0.102	0.40
	Modified Spring	616	0.146	0.41
	Modified Fall	788	0.118	0.39
Axis 4	Baseline	895	0.041	0.41
	Modified Spring	614	0.080	0.43
	Modified Fall	787	0.056	0.43
Axis 5	Baseline	882	0.036	0.39
	Modified Spring	611	0.086	0.39
	Modified Fall	776	0.050	0.39
Overall	Baseline	906	0.005	0.18
	Modified Spring	624	0.035	0.18
	Modified Fall	803	0.004	0.16

Overall modified fall results show effectively no change with respect to the baseline course (Table D.4). Axis two, five, and overall still exhibit significant epistemic deterioration in students. These students came into the fall with more well-developed beliefs than their spring (and baseline) counterparts, but inevitably still ended up with a decline in EBAPS performance.

Table D.4: Modified Fall Pre-Test and Post-Test results by EBAPS axis. Scores range from 0 to 4. Bold indicates prominence with $|d| \geq .15$ and $p \leq .05$. Axis 1: Structure of scientific knowledge, Axis 2: Nature of knowing and learning, Axis 3: Real-life applicability, Axis 4: Evolving knowledge, Axis 5: Source of ability to learn

	Pre-Test	s.d.	Post-Test	s.d.	effect size	p value
Axis 1	2.32	0.46	2.32	0.48	0.01	0.72
Axis 2	2.62	0.50	2.54	0.53	-0.16	0.00
Axis 3	2.93	0.70	2.92	0.69	-0.01	0.69
Axis 4	2.69	0.78	2.58	0.77	-0.14	0.00
Axis 5	3.23	0.60	3.01	0.70	-0.33	0.00
Overall	2.69	0.35	2.62	0.39	-0.19	0.00

Table D.5 shows results for the modified fall students by College. Students here again follow effectively similar negative decays as found in the baseline. Along axes where we have noticeable decays ($d \leq -0.15$ with $p < 0.05$) most effect sizes are around or above $|d| = 0.25$. Much like in the baseline data numerous other axes are on the cusp of showing noticeable decline as well. Unlike the baseline, axis one has now undergone a prominent decline for ED students as well as the overall axis for LS students. No visible decays for UC students are seen in the modified fall data, however.

Table D.6 shows results for the modified fall students by Degree. Modified fall BS students maintain a similar evolutionary profile to those in the baseline. Fall modified BA students experience decays along axes two, five, and overall. Decays along these axes were hinted at in the baseline students but display a clear signal in modified fall data, even with a smaller sample size.

Table D.7 shows results for the modified fall students by Gender. Fall modified data yields more prominent occurrences of epistemic decay amongst females, but is consistent with the baseline in the finding that males struggle along axis four in comparison. The course changes have consistently shown to do effectively nothing to the students in the fall portion of the modified course.

Table D.8 shows results for the modified fall students by Status. Modified fall data show strong negative prominences across multiple axes and statuses. Of noticeable difference from the baseline are the SR students, who see detectable decays

Table D.5: EBAPS Pre-Test and Post-Test results for Fall Modified by College. Scores range from 0 to 4. Bold indicates prominence with $|d| \geq .15$ and $p \leq .05$. Axis 1: Structure of scientific knowledge, Axis 2: Nature of knowing and learning, Axis 3: Real-life applicability, Axis 4: Evolving knowledge, Axis 5: Source of ability to learn

College	Axis	Pre-Test	s.d.	Post-Test	s.d.	effect size	p value	N
AA	Axis 1	2.38	0.41	2.37	0.51	-0.02	0.65	105
	Axis 2	2.60	0.50	2.50	0.49	-0.21	0.05	
	Axis 3	2.95	0.66	2.94	0.66	-0.01	0.91	
	Axis 4	2.79	0.74	2.60	0.81	-0.24	0.10	
	Axis 5	3.21	0.70	2.98	0.75	-0.31	0.00	
	Overall	2.71	0.37	2.62	0.38	-0.24	0.00	
BU	Axis 1	2.19	0.42	2.21	0.42	0.05	0.35	222
	Axis 2	2.58	0.50	2.49	0.52	-0.17	0.01	
	Axis 3	2.79	0.64	2.77	0.68	-0.03	0.91	
	Axis 4	2.56	0.81	2.45	0.76	-0.14	0.09	
	Axis 5	3.25	0.58	2.94	0.71	-0.47	0.00	
	Overall	2.62	0.31	2.54	0.36	-0.25	0.00	
ED	**Axis 1**	2.32	0.46	2.22	0.44	-0.23	0.05	88
	Axis 2	2.64	0.46	2.55	0.53	-0.19	0.14	
	Axis 3	2.91	0.64	2.79	0.64	-0.18	0.22	
	Axis 4	2.76	0.71	2.66	0.75	-0.14	0.43	
	Axis 5	3.14	0.68	3.02	0.61	-0.19	0.07	
	Overall	2.69	0.35	2.59	0.36	-0.27	0.02	
EN	Axis 1	2.46	0.51	2.52	0.47	0.11	0.21	97
	Axis 2	2.70	0.52	2.60	0.61	-0.18	0.25	
	Axis 3	3.17	0.74	3.22	0.75	0.07	0.64	
	Axis 4	2.70	0.79	2.63	0.70	-0.10	0.38	
	Axis 5	3.33	0.57	3.17	0.71	-0.25	0.02	
	Overall	2.80	0.34	2.75	0.42	-0.14	0.43	
LS	Axis 1	2.38	0.48	2.39	0.48	0.02	0.96	167
	Axis 2	2.68	0.51	2.61	0.53	-0.14	0.12	
	Axis 3	3.05	0.70	2.97	0.71	-0.11	0.09	
	Axis 4	2.75	0.75	2.62	0.78	-0.17	0.20	
	Axis 5	3.23	0.59	2.97	0.74	-0.39	0.00	
	Overall	2.74	0.36	2.67	0.40	-0.21	0.00	
UC	Axis 1	2.28	0.47	2.29	0.50	0.02	0.60	110
	Axis 2	2.57	0.47	2.50	0.52	-0.14	0.24	
	Axis 3	2.80	0.80	2.90	0.63	0.13	0.23	
	Axis 4	2.65	0.79	2.58	0.78	-0.08	0.67	
	Axis 5	3.17	0.53	3.09	0.64	-0.15	0.25	
	Overall	2.64	0.36	2.62	0.39	-0.05	0.92	

along axes two, five, and overall. Epistemic increases along axis one for sophomore students are no longer occurring and junior students now actually undergo prominent decreases in axis two performance.

Table D.6: EBAPS Modified Fall Pre-Test and Post-Test results by Degree. Scores range from 0 to 4. Bold indicates prominence with $|d| \geq .15$ and $p \leq .05$. Axis 1: Structure of scientific knowledge, Axis 2: Nature of knowing and learning, Axis 3: Real-life applicability, Axis 4: Evolving knowledge, Axis 5: Source of ability to learn

Degree	Axis	Pre-Test	s.d.	Post-Test	s.d.	effect size	p value	N
BA	Axis 1	2.38	0.45	2.39	0.49	0.02	0.85	152
	Axis 2	2.64	0.49	2.52	0.49	-0.23	0.01	
	Axis 3	2.98	0.68	2.96	0.67	-0.03	0.57	
	Axis 4	2.75	0.73	2.67	0.81	-0.10	0.46	
	Axis 5	3.20	0.66	2.98	0.73	-0.31	0.00	
	Overall	2.72	0.36	2.65	0.38	-0.19	0.00	
BS	Axis 1	2.30	0.47	2.30	0.47	0.01	0.55	627
	Axis 2	2.62	0.50	2.54	0.54	-0.14	0.00	
	Axis 3	2.91	0.71	2.90	0.70	-0.02	0.75	
	Axis 4	2.66	0.79	2.55	0.76	-0.14	0.01	
	Axis 5	3.23	0.58	3.02	0.69	-0.34	0.00	
	Overall	2.68	0.35	2.61	0.39	-0.18	0.00	

Table D.7: EBAPS Modified Fall Pre-Test and Post-Test results by Gender. Scores range from 0 to 4. Bold indicates prominence with $|d| \geq .15$ and $p \leq .05$. Axis 1: Structure of scientific knowledge, Axis 2: Nature of knowing and learning, Axis 3: Real-life applicability, Axis 4: Evolving knowledge, Axis 5: Source of ability to learn

Gender	Axis	Pre-Test	s.d.	Post-Test	s.d.	effect size	p value	N
Female	Axis 1	2.33	0.44	2.33	0.47	0.01	0.78	352
	Axis 2	2.68	0.48	2.60	0.51	-0.16	0.01	
	Axis 3	2.89	0.66	2.92	0.66	0.03	0.50	
	Axis 4	2.72	0.76	2.66	0.76	-0.09	0.30	
	Axis 5	3.17	0.65	3.00	0.70	-0.26	0.00	
	Overall	2.70	0.35	2.64	0.38	-0.16	0.01	
Male	Axis 1	2.33	0.49	2.33	0.48	0.00	0.97	349
	Axis 2	2.59	0.50	2.51	0.53	-0.16	0.01	
	Axis 3	2.98	0.74	2.93	0.71	-0.06	0.24	
	Axis 4	2.63	0.78	2.52	0.77	-0.15	0.04	
	Axis 5	3.29	0.56	3.03	0.69	-0.40	0.00	
	Overall	2.70	0.35	2.62	0.39	-0.22	0.00	

Table D.8: EBAPS Modified Fall Pre-Test and Post-Test results by Status. Scores range from 0 to 4. Bold indicates prominence with $|d| \geq .15$ and $p \leq .05$. Axis 1: Structure of scientific knowledge, Axis 2: Nature of knowing and learning, Axis 3: Real-life applicability, Axis 4: Evolving knowledge, Axis 5: Source of ability to learn

Status	Axis	Pre-Test	s.d.	Post-Test	s.d.	effect size	p value	N
FR	Axis 1	2.26	0.45	2.28	0.44	0.05	0.26	446
	Axis 2	2.58	0.49	2.51	0.51	-0.14	0.01	
	Axis 3	2.87	0.70	2.81	0.67	-0.08	0.12	
	Axis 4	2.62	0.79	2.56	0.80	-0.08	0.25	
	Axis 5	3.25	0.58	3.01	0.71	-0.37	0.00	
	Overall	2.66	0.33	2.59	0.37	-0.19	0.00	
SO	Axis 1	2.38	0.47	2.33	0.50	-0.09	0.27	206
	Axis 2	2.60	0.49	2.56	0.54	-0.07	0.41	
	Axis 3	2.96	0.69	3.02	0.69	0.08	0.27	
	Axis 4	2.72	0.71	2.61	0.73	-0.15	0.16	
	Axis 5	3.18	0.62	3.07	0.66	-0.16	0.03	
	Overall	2.69	0.36	2.65	0.38	-0.11	0.12	
FR	Axis 1	2.36	0.44	2.35	0.48	-0.01	0.91	81
	Axis 2	2.76	0.46	2.58	0.59	-0.34	0.01	
	Axis 3	2.96	0.69	2.99	0.66	0.04	0.83	
	Axis 4	2.89	0.80	2.60	0.72	-0.39	0.01	
	Axis 5	3.17	0.62	2.86	0.72	-0.45	0.00	
	Overall	2.75	0.34	2.62	0.40	-0.34	0.00	
SR	Axis 1	2.50	0.51	2.53	0.55	0.05	0.62	71
	Axis 2	2.80	0.55	2.63	0.60	-0.29	0.04	
	Axis 3	3.14	0.68	3.19	0.79	0.07	0.61	
	Axis 4	2.73	0.78	2.56	0.75	-0.22	0.15	
	Axis 5	3.29	0.62	3.02	0.71	-0.40	0.00	
	Overall	2.85	0.37	2.74	0.47	-0.25	0.04	

Chi-Squared Tables Item Analysis

Table D.9: Pretest Proportional Information for EBAPS Question 2. Subscripts with identical letters represent proportions which do not significantly differ at the $p < 0.05$ level

Question 2				
Response Type	Statistic	Baseline	Modified Spring	Modified Fall
0 (Unsophisticated)	Count	$754_{a,b}$	498_{b}	692_{a}
	Expected Count	754.6	519.7	669.7
	% Within	83.2%	79.8%	86.1%
2 (Sophisticated)	Count	$152_{a,b}$	126_{b}	112_{a}
	Expected Count	151.4	104.3	134.3
	% Within	16.8%	20.2%	13.9%

Table D.10: Pre-test Proportional Information for EBAPS Question 4. Subscripts with identical letters represent proportions which do not significantly differ at the $p < 0.05$ level

Question 4				
Response Type	Statistic	Baseline	Modified Spring	Modified Fall
0 (Unsophisticated)	Count	137_{a}	88_{a}	120_{a}
	Expected Count	133.9	92.2	118.8
	% Within	15.1%	14.1%	14.9%
1 (Neutral)	Count	$227_{a,b}$	170_{b}	166_{a}
	Expected Count	218.5	150.5	193.9
	% Within	25.1%	27.2%	20.6%
2 (Sophisticated)	Count	542_{a}	366_{a}	518_{a}
	Expected Count	553.5	381.2	491.2
	% Within	59.8%	58.7%	64.4%

Table D.11: Pretest Proportional Information for EBAPS Question 5. Subscripts with identical letters represent proportions which do not significantly differ at the $p < 0.05$ level

Question 5				
Response Type	Statistic	Baseline	Modified Spring	Modified Fall
0 (Unsophisticated)	Count	67_a	47_a	53_a
	Expected Count	64.8	44.6	57.5
	% Within	7.4%	7.5%	6.6%
1 (Neutral)	Count	68_a	$38_{a,b}$	31_b
	Expected Count	52.4	36.1	46.5
	% Within	7.5%	5.8%	3.9%
2 (Sophisticated)	Count	771_a	$541_{a,b}$	720_a
	Expected Count	788.8	543.3	700.0
	% Within	85.1%	86.7%	89.6%

Table D.12: Pretest Proportional Information for EBAPS Question 6. Subscripts with identical letters represent proportions which do not significantly differ at the $p < 0.05$ level

Question 6				
Response Type	Statistic	Baseline	Modified Spring	Modified Fall
0 (Unsophisticated)	Count	135_a	61_b	$97_{a,b}$
	Expected Count	113.7	78.3	100.9
	% Within	14.9%	9.8%	12.1%
1 (Neutral)	Count	202_a	145_a	170_a
	Expected Count	200.7	138.2	178.1
	% Within	22.3%	23.2%	21.1%
2 (Sophisticated)	Count	569_a	418_a	537_a
	Expected Count	591.6	407.4	525.0
	% Within	62.8%	67.0%	66.8%

Table D.13: Pretest Proportional Information for EBAPS Question 8. Subscripts with identical letters represent proportions which do not significantly differ at the $p < 0.05$ level

Question 8				
Response Type	Statistic	Baseline	Modified Spring	Modified Fall
0 (Unsophisticated)	Count	196_a	136_a	119_b
	Expected Count	175.1	120.6	155.4
	% Within	21.6%	21.8%	14.8%
2 (Sophisticated)	Count	710_a	488_a	685_b
	Expected Count	730.9	503.4	648.6
	% Within	78.4%	78.2%	85.2%

Table D.14: Pretest Proportional Information for EBAPS Question 9. Subscripts with identical letters represent proportions which do not significantly differ at the $p < 0.05$ level

Question 9				
Response Type	Statistic	Baseline	Modified Spring	Modified Fall
0 (Unsophisticated)	Count	108_a	72_a	53_b
	Expected Count	90.4	62.3	80.3
	% Within	11.9%	11.5%	6.6%
1 (Neutral)	Count	120_a	88_a	82_a
	Expected Count	112.6	77.5	99.9
	% Within	13.2%	14.1%	10.2%
2 (Sophisticated)	Count	678_a	464_a	669_b
	Expected Count	703.0	484.2	623.8
	% Within	74.8%	74.4%	83.2%

Table D.15: Pretest Proportional Information for EBAPS Question 11. Subscripts with identical letters represent proportions which do not significantly differ at the $p < 0.05$ level

Question 11				
Response Type	Statistic	Baseline	Modified Spring	Modified Fall
0 (Unsophisticated)	Count	129_a	89_a	85_a
	Expected Count	117.6	81.0	104.4
	% Within	14.2%	14.3%	10.6%
1 (Neutral)	Count	175_a	166_b	201_b
	Expected Count	210.4	144.9	186.7
	% Within	19.3%	26.6%	25.0%
2 (Sophisticated)	Count	602_a	369_b	$518_{a,b}$
	Expected Count	578.0	398.1	512.9
	% Within	66.4%	59.1%	64.4%

Table D.16: Pretest Proportional Information for EBAPS Question 16. Subscripts with identical letters represent proportions which do not significantly differ at the $p < 0.05$ level

Question 16				
Response Type	Statistic	Baseline	Modified Spring	Modified Fall
0 (Unsophisticated)	Count	94_a	69_a	54_b
	Expected Count	84.2	58.0	74.8
	% Within	10.4%	11.1%	6.7%
1 (Neutral)	Count	90_a	62_a	57_a
	Expected Count	81.1	55.9	72.0
	% Within	9.9%	9.9%	7.1%
2 (Sophisticated)	Count	722_a	493_a	693_b
	Expected Count	740.6	510.1	657.3
	% Within	79.7%	79.0%	86.2%

Table D.17: Pretest Proportional Information for EBAPS Question 17. Subscripts with identical letters represent proportions which do not significantly differ at the $p < 0.05$ level

Question 17				
Response Type	Statistic	Baseline	Modified Spring	Modified Fall
0 (Unsophisticated)	Count	599_a	351_b	$499_{a,b}$
	Expected Count	562.5	387.4	499.1
	% Within	66.1%	56.3%	62.1%
2 (Sophisticated)	Count	307_a	273_b	$305_{a,b}$
	Expected Count	343.5	236.6	304.9
	% Within	33.9%	43.8%	37.9%

Table D.18: Pre-test Proportional Information for EBAPS Question 21. Subscripts with identical letters represent proportions which do not significantly differ at the $p < 0.05$ level

Question 21				
Response Type	Statistic	Baseline	Modified Spring	Modified Fall
0 (Unsophisticated)	Count	21_a	20_a	14_a
	Expected Count	21.3	14.7	18.9
	% Within	2.3%	3.2%	1.7%
1 (Neutral)	Count	201_a	128_a	142_a
	Expected Count	182.8	124.9	162.2
	% Within	22.2%	20.5%	17.7%
2 (Sophisticated)	Count	684_a	$476_{a,b}$	648_b
	Expected Count	701.8	483.4	622.8
	% Within	75.5%	76.3%	80.6%

Table D.19: Pretest Proportional Information for EBAPS Question 22. Subscripts with identical letters represent proportions which do not significantly differ at the $p < 0.05$ level

Question 22				
Response Type	Statistic	Baseline	Modified Spring	Modified Fall
0 (Unsophisticated)	Count	56_a	$25_{a,b}$	19_b
	Expected Count	38.8	26.7	34.4
	% Within	6.2%	4.0%	2.4%
1 (Neutral)	Count	192_a	129_a	136_a
	Expected Count	177.4	122.2	157.4
	% Within	21.2%	20.7%	16.9%
2 (Sophisticated)	Count	658_a	470_a	649_b
	Expected Count	689.8	475.1	612.1
	% Within	72.6%	75.3%	80.7%

Table D.20: Pretest Proportional Information for EBAPS Question 25. Subscripts with identical letters represent proportions which do not significantly differ at the $p < 0.05$ level

Question 25				
Response Type	Statistic	Baseline	Modified Spring	Modified Fall
0 (Unsophisticated)	Count	239_a	148_a	122_b
	Expected Count	197.6	136.1	175.3
	% Within	26.4%	23.7%	15.2%
1 (Neutral)	Count	144_a	96_a	140_a
	Expected Count	147.5	101.6	130.9
	% Within	15.9%	15.4%	17.4%
2 (Sophisticated)	Count	523_a	380_a	542_b
	Expected Count	560.9	386.3	497.8
	% Within	57.7%	60.9%	67.4%

Table D.21: Pretest Proportional Information for EBAPS Question 28. Subscripts with identical letters represent proportions which do not significantly differ at the $p < 0.05$ level

Question 28				
Response Type	Statistic	Baseline	Modified Spring	Modified Fall
0 (Unsophisticated)	Count	112_a	88_a	81_a
	Expected Count	109.1	75.1	96.8
	% Within	12.4%	14.1%	10.1%
1 (Neutral)	Count	148_a	103_a	102_a
	Expected Count	137.0	94.4	121.6
	% Within	16.3%	16.5%	12.7%
2 (Sophisticated)	Count	646_a	433_a	621_b
	Expected Count	659.9	454.5	585.6
	% Within	71.3%	69.4%	77.2%

Table D.22: Post-test Proportional Information for EBAPS Question 3. Subscripts with identical letters represent proportions which do not significantly differ at the $p < 0.05$ level

Question 3				
Response Type	Statistic	Baseline	Modified Spring	Modified Fall
0 (Unsophisticated)	Count	178_a	$108_{a,b}$	110_b
	Expected Count	153.7	105.9	136.4
	% Within	19.6%	17.3%	13.7%
1 (Neutral)	Count	180_a	113_a	155_a
	Expected Count	173.9	119.8	154.3
	% Within	19.9%	18.1%	19.3%
2 (Sophisticated)	Count	548_a	$403_{a,b}$	539_b
	Expected Count	578.4	398.4	513.3
	% Within	60.5%	64.6%	67.0%

Table D.23: Post-test Proportional Information for EBAPS Question 6. Subscripts with identical letters represent proportions which do not significantly differ at the $p < 0.05$ level

Question 6				
Response Type	Statistic	Baseline	Modified Spring	Modified Fall
0 (Unsophisticated)	Count	137_a	76_a	98_a
	Expected Count	120.7	83.1	107.1
	% Within	15.1%	12.2%	12.2%
1 (Neutral)	Count	240_a	129_b	$181_{a,b}$
	Expected Count	213.5	147.0	189.5
	% Within	26.5%	20.7%	22.5%
2 (Sophisticated)	Count	529_a	419_b	525_b
	Expected Count	571.8	393.9	507.4
	% Within	58.4%	67.1%	65.3%

Table D.24: Post-test Proportional Information for EBAPS Question 8. Subscripts with identical letters represent proportions which do not significantly differ at the $p < 0.05$ level

Question 8				
Response Type	Statistic	Baseline	Modified Spring	Modified Fall
0 (Unsophisticated)	Count	229_a	124_b	$165_{a,b}$
	Expected Count	201.1	138.5	178.4
	% Within	25.3%	19.9%	20.5%
2 (Sophisticated)	Count	677_a	500_b	$639_{a,b}$
	Expected Count	704.9	485.5	625.5
	% Within	74.7%	80.1%	79.5%

Table D.25: Post-test Proportional Information for EBAPS Question 9. Subscripts with identical letters represent proportions which do not significantly differ at the $p < 0.05$ level

Question 9				
Response Type	Statistic	Baseline	Modified Spring	Modified Fall
0 (Unsophisticated)	Count	165_a	105_a	91_b
	Expected Count	140.1	96.5	124.4
	% Within	18.2%	16.8%	11.3%
1 (Neutral)	Count	139_a	78_a	122_a
	Expected Count	131.6	90.6	116.8
	% Within	15.3%	12.5%	15.2%
2 (Sophisticated)	Count	602_a	$441_{a,b}$	591_b
	Expected Count	634.3	436.9	562.9
	% Within	66.4%	70.7%	73.5%

Table D.26: Post-test Proportional Information for EBAPS Question 14. Subscripts with identical letters represent proportions which do not significantly differ at the $p < 0.05$ level

Question 14				
Response Type	Statistic	Baseline	Modified Spring	Modified Fall
0 (Unsophisticated)	Count	123_a	73_a	80_a
	Expected Count	107.1	73.8	95.1
	% Within	13.6%	11.7%	10.0%
1 (Neutral)	Count	133_a	68_a	106_a
	Expected Count	119.2	82.1	105.8
	% Within	14.7%	10.9%	13.2%
2 (Sophisticated)	Count	650_a	484_b	618_b
	Expected Count	679.7	468.1	603.2
	% Within	71.7%	77.4%	76.9%

Table D.27: Post-test Proportional Information for EBAPS Question 17. Subscripts with identical letters represent proportions which do not significantly differ at the $p < 0.05$ level

Question 17				
Response Type	Statistic	Baseline	Modified Spring	Modified Fall
0 (Unsophisticated)	Count	565_a	335_b	448_b
	Expected Count	523.3	360.4	464.3
	% Within	62.4%	53.7%	55.7%
2 (Sophisticated)	Count	341_a	289_b	356_b
	Expected Count	382.7	263.6	339.7
	% Within	37.6%	46.3%	44.3%

Table D.28: Post-test Proportional Information for EBAPS Question 21. Subscripts with identical letters represent proportions which do not significantly differ at the $p < 0.05$ level

Question 21				
Response Type	Statistic	Baseline	Modified Spring	Modified Fall
0 (Unsophisticated)	Count	48_a	$22_{a,b}$	17_b
	Expected Count	33.8	23.3	30.0
	% Within	5.3%	3.5%	2.1%
1 (Neutral)	Count	215_a	$122_{a,b}$	133_b
	Expected Count	182.4	125.7	161.9
	% Within	23.7%	19.6%	16.5%
2 (Sophisticated)	Count	643_a	480_b	654_b
	Expected Count	689.8	475.1	612.1
	% Within	71.0%	76.9%	81.3%

Table D.29: Post-test Proportional Information for EBAPS Question 22. Subscripts with identical letters represent proportions which do not significantly differ at the $p < 0.05$ level

Question 22				
Response Type	Statistic	Baseline	Modified Spring	Modified Fall
0 (Unsophisticated)	Count	88_a	$47_{a,b}$	51_b
	Expected Count	72.2	49.7	64.1
	% Within	9.7%	7.5%	6.3%
1 (Neutral)	Count	235_a	$129_{a,b}$	154_b
	Expected Count	201.1	138.5	178.4
	% Within	25.9%	20.7%	19.2%
2 (Sophisticated)	Count	583_a	448_b	599_b
	Expected Count	632.7	435.8	561.5
	% Within	64.3%	71.8%	74.5%

Table D.30: Post-test Proportional Information for EBAPS Question 24. Subscripts with identical letters represent proportions which do not significantly differ at the $p < 0.05$ level

Question 24				
Response Type	Statistic	Baseline	Modified Spring	Modified Fall
0 (Unsophisticated)	Count	200_a	141_a	170_a
	Expected Count	198.4	136.6	176.0
	% Within	22.1%	22.6%	21.1%
1 (Neutral)	Count	149_a	102_a	95_b
	Expected Count	134.3	92.5	119.2
	% Within	16.4%	16.3%	11.8%
2 (Sophisticated)	Count	557_a	381_a	539_a
	Expected Count	573.3	394.9	508.8
	% Within	61.5%	61.1%	67.0%

Table D.31: Post-test Proportional Information for EBAPS Question 25. Subscripts with identical letters represent proportions which do not significantly differ at the $p < 0.05$ level

Question 25				
Response Type	Statistic	Baseline	Modified Spring	Modified Fall
0 (Unsophisticated)	Count	274_a	$161_{a,b}$	190_b
	Expected Count	242.6	167.1	215.3
	% Within	30.2%	25.8%	23.6%
1 (Neutral)	Count	198_a	123_a	146_a
	Expected Count	181.3	124.9	160.9
	% Within	21.9%	19.7%	18.2%
2 (Sophisticated)	Count	434_a	340_b	468_b
	Expected Count	482.1	332.1	427.8
	% Within	47.9%	54.5%	58.2%

Table D.32: Post-test Proportional Information for EBAPS Question 30. Subscripts with identical letters represent proportions which do not significantly differ at the $p < 0.05$ level

Question 30				
Response Type	Statistic	Baseline	Modified Spring	Modified Fall
0 (Unsophisticated)	Count	195_a	125_a	150_a
	Expected Count	182.4	125.7	161.9
	% Within	21.5%	20.0%	18.7%
1 (Neutral)	Count	$174_{a,b}$	128_b	121_a
	Expected Count	164.2	113.1	145.7
	% Within	19.2%	20.5%	15.0%
2 (Sophisticated)	Count	537_a	371_a	533_b
	Expected Count	559.4	385.3	496.4
	% Within	59.3%	59.5%	66.3%

APPENDIX E

MATERIAL USED WITHIN MODIFIED COURSE

Science in the Media

Read an online article from any .edu or .gov source. The article should be about science, but does not need to be about astronomy in particular.
Source:
Topic:
Summary of content
Comments related to source
Is it reliable?
Are personal beliefs driving the claim?
Do you agree with the conclusions drawn by the author? Why or why not?
Can you find a website/source that makes the opposite claim?
Link/source:
Is this site reputable? How did you decide this?
Do other sources agree with this source? Most, some or none?
Where does the preponderance of evidence point?

Exit Tickets

Before beginning the exam review (which typically occurs two days before the exam) we ask students to jot down one concept they are still struggling with, known as an exit ticket or the muddiest point. [17] Using the course learning goals as their guide, students think about what they do and do not understand and why, thus engaging in metacognition (axis two), which can help focus their studies. These slips of paper are collected and read by the instructors so we can glean insight into what students struggle with most. Answers are collated and inform teaching for the next semester. A second technique we employ occurs during review sessions, when students are asked to write their own exam questions. Research in a psychology course [82] found that students who wrote their own test questions (either multiple choice or essay style) performed better on two out of three tests compared to students who wrote no questions. Even if the quality of the questions is not high, it appears this is an effective study strategy for students. We also collect and read these.

Exam Wrappers

After each of the first two midterms we ask students to complete an exam wrapper. [141] In this reflective writing assignment, students discuss how they prepared for the exam - alone or with a group, creating flash cards or going over course materials. Students are then asked to go over their exam and look for any patterns in the questions they missed. Was one topic not understood? Did they perform poorly just on the multiple choice or just on the short answer? Finally, students are asked how they might change their study habits for the next exam. The goal of these exercises is to again have students engage in metacognition (axis two) to improve their understanding before the exam, then reflect on their exam preparation and performance after their exams have been graded and handed back.

Visual Representations of Knowledge

Finally, as a regular part of the course, students are asked to create visual representations of their understanding, including the following examples. By drawing out a pictoral representation of an in-class demonstration, students will process information differently than by just taking notes [145, 152] and hence must practice metacognitive strategies (axis two) to translate between their notes and the representation they are creating.

A demonstration of seasonal asterisms at the front of the classroom utilized an Earth globe, around which several students acted as stars. Following this, students are asked to draw a picture in their notes that shows why Leo is a seasonal asterism. During the lecture on stellar evolution, students are asked to create a concept map describing how stars evolve and showing how that the evolutionary sequence depends on the mass of the star. After taking down this information in notes, students can create a visual representation of the entire process and sum up their knowledge in a visual way. Finally, students are asked to create a Venn diagram comparing Jovian and Terrestrial planets during the section on the solar system. Using class time to let students sift through information is one way we demonstrate to our students that tasks related to metacognition are worthwhile and can increase conceptual understanding.

APPENDIX F

WRITTEN RESPONSES

Written Responses

Science

The question as posed to students was:

> "What is your definition of science?"

The question is straightforward, intending to yield insight into what students view science to be.

There were 79 total written responses in which students provided their definition of science. Two very noticeable themes appeared within the responses, the most frequent being the view that science is analyzing. Within this theme, science was either the analysis of natural/physical things:

> "I define the word science as the study of natural and physical things with studies and experiments."

> "Science is an activity that involves intellectual activity, by studying and observing the natural world."

or science was the analysis of life/surroundings:

> "My definition of science is the study of our surroundings and how they work."

> "My definition of science is the study of the physical environment surrounding living beings as well as the interaction between living and non-living things in order to survive."

The next most frequently occurring theme was the treatment of science as being a type of methodological process:

> "Science is the study of hypothesizing things about the natural and physical world, and going forth and testing these hypotheses to create theories about the way our world works."

> "Studying the natural world in a systematic way through observation and experimentation."

> "A way of interacting with the world that includes observing, asking questions, and testing theories."

The students who focused on the process of science intrinsically also linked it to being either the analysis of natural/physical things or the analysis of life/surroundings. Few students treated science as purely being a method for testing, as seen in the example below:

"The sequential steps beginning with observation, testing the observation, and recording results."

Other themes beyond the two above were seen within the responses, but not with as great of frequency. One such category was the treatment of science as fact/proof, a rather unsophisticated view:

"My definition of science is something like using evidence or physical proof to disprove or prove that something exists."

"My definition of science is any practice that uses facts to define new theories about the way that things work."

Similar to the treatment of science as a process is a theme which viewed science as being the uncovering of information through reasoning and/or logic:

"My definition of science is using analytical reasoning to better understand the world around us."

"My definition of science is a mathematical and logical inquiry to the physical aspects of the world and universe around us."

The last theme to be discussed represents an all-encompassing view from the students that science is anything/everything:

"Science is everything we study that can be tested and observed."

"Science is something that humans have had no effect on. Everything that was created during the big bang is science. Everything is science."

Considering the findings present above, a general definition of science from students would be that science is the exploration of our surroundings in an attempt to understand the how/why of an observation, as it pertains to the physical world. Those individuals who focus more on the process of science, the systematic approach involved in the study of the physical world are considered among the most sophisticated responses.

Natural Ability

The question as posed to students was:

> "Consider the following statement: 'To be successful in science, hard work and natural ability are factors that may play a role.' Define what you believe hard work is in this context. Also define what you believe natural ability is in this context."

The question is intended to probe what students interpret both "hard work" and "natural ability" to be, within the context of being successful in science. In this section, student definitions of "natural ability" are discussed.

There were 78 total written responses which clearly portrayed a definition of natural ability. There were two prominent themes within the responses, the first associating with a more traditional definition:

> "Natural ability would be God given affinities for science."

> "Natural ability to me means the natural born gift to understand and determine science in a better understanding."

> "Natural ability is what you are born with the ability to do and you're not taught it."

These students associate natural ability with some form of intrinsic, god-given ability. A trait which makes a student inherently more capable of understanding science. Some students go so far as to attribute this to intelligence:

> "Natural ability helps if you are better at remembering things."

> "I believe 'natural ability' in physics is, perhaps, a combination of mathematical aptitude, pattern-recognition ability and an ability to solve complex, novel problems."

while others attribute it to intuitive ability, instincts:

> "Natural ability might be a kind of intuition or some kind of sense that makes some better at scientific thought and reasoning than others."

> "I believe natural ability is more along the lines of using your senses and skills and applying them to the specific context."

Most students simply associate it with no specific ability other than one which makes someone "better" at science. The next most prominent theme was the belief that natural ability simply makes the material easier to understand/learn, that you have to do less work than others:

> "Natural abilities would be skills that you pick up with relative ease."

> "Your natural ability is how easy it is for you to develop a new skill or preform the one at hand."

> "Natural ability might be to naturally have a grasp on science without working too hard."

These students make no mention of any type of god-given abilities, but rather just state that natural ability refers to how easy it is for a student to understand and/or learn the material. Although one may expect that these students are refering to natural born ability, further categories show that this should not be assumed. A separate, but potential subset, of the category above is a belief in what may be called "quick learning":

> "Natural ability in this statement is just understanding material easily and being able to learn it in a quick and efficient manner."

> "[. . .] you are able to understand it quickly and apply it to problems right away. You might understand a unit faster than your peers."

The crux of their view is that those who have natural ability in something will learn something quickly. Students here were also seen to imply that hard work, as discussed further below, may not be necessary. The next most commonly occurring theme was focused on intrinsic interest:

> "I think 'natural ability' means an interest in the subject, it is easier to learn and continue with a topic if you enjoy the topic."

> "Natural ability is the spark that helps people be interested in a scientific field. It makes solving things easier and new concepts are simpler to understand."

> "Natural ability to me would be actually wanting to do the work basically it's second nature for you. Science has to be fun to you rather than it giving you headaches and disappointment. You indeed have to have a passion for it."

Students here actually attribute natural ability to nothing more than intrinsic interest in the subject. Whether this interest was nature or nurture is unclear. The last notable theme present within the student written responses is the emphasis on prior knowledge:

> "Natural ability is understanding certain concepts that you just kinda know about because of something you learned early on. Things you do not need to think hard about and you just know them already."

> "Natural ability would be an interest in what you are studying and a general knowledge about some of the information. Natural ability would also consist of being able to inform others easily on the subject and being able to understand and learn the material with ease."

Students here attribute natural ability as being nothing more than prior knowledge that a student has before tackling the task at hand. The latter response above actually connects this belief in with both interest for the subject and with the ability to understand the material with ease.

Hard Work

The question as posed to students was:

> "Consider the following statement: 'To be successful in science, hard work and natural ability are factors that may play a role.' Define what you believe hard work is in this context. Also define what you believe natural ability is in this context."

The question is intended to probe what students interpret both "hard work" and "natural ability" to be, within the context of being successful in science. In this section, student definitions of "hard work" are discussed.

There were 74 total written responses which portrayed a clear definition of hard work. Much like natural ability, two prominent themes were observed within the responses. The first prominent theme was the belief that hard work is making sure you understand the material:

> "Hard work is making effort to fully grasp the concepts."

> "Hard work is the combination of practicing formulas and solving problems, with an understanding of the reasoning behind it."

> "Hard work is studying hard and applying yourself more in order to understand fully what you're inspecting."

As can be seen from the student examples above, hard work is viewed as simply working to attain an understanding of the material. Unfortunately, it is not clear what is meant by "understanding", an issue which also shows up within the EBAPS question responses. The second of the above responses indicates that "understanding" (to that individual) implies knowing why particular formulas/methods are employed. The other prominent theme present within the response data was dedication to learning:

> "I believe hard work is the equivalent of diligence. Consistent work, for long periods of time, is hard work."

> "Working hard is putting in the time and effort to excel at a subject."

> "Hard work, definitely, is putting in the time to study and be interested in your field."

The most general aspect of these responses likened hard work to time and effort put into learning about the subject. Some students, such as the latter student in the above examples, linked hard work with an interest or passion for the subject. There exist smaller, possible subsets of both the prominent themes discussed above. One such subset was an emphasis on hard work as being prepared:

> "I would describe hard work to be preparing for the lectures before hand (reading over lecture notes for the day as well as the corresponding chapter in the textbook) [. . .]."

Another was hard work as being studying/practicing the material:

> "I believe that in this statement 'hard work' means applying yourself to the study and practicing and perfecting it every day."

Lastly, there was a rather unsophisticated subset which viewed hard work as memorization:

> "Hard work is being able to memorize multiple equations and facts about science."

The remaining themes were all interconnected, but frequent in their occurrences. These themes were perseverance:

> "To me hard work means that someone continually puts effort into a situation and doesn't give up when they are annoyed or confused."

perseverance and dedication:

> "Hard work in this context means to always study and to always keep working as hard as you can to insure you are moving forward and learning something new every day."

and finally, dedication/perseverance and understanding:

> "Hard work is dedicating yourself to fully understanding concepts, using your own time, and actively participating."

Within these responses, perseverance refers to continuing to work with the material despite setbacks, dedication refers to relentlessly working with the material, and understanding is similar to the theme above in which students place emphasis on "understanding" the material.

When considering all themes, students generally believe hard work to be committing to learning/understanding the material, regardless of the time, effort, or setbacks involved.

<u>EBAPS Question 1</u>

The question as posed to students was:

> "You read something in your science textbook that seems to disagree with your own experiences. But to learn science well, you shouldn't think about your own experiences; you should just focus on what the book says. Do you agree or disagree with this viewpoint? Explain your thoughts."

The question is intended to be representative of EBAPS question 1.

There are several concerns about this question pertaining to the manner in which students responded. The first concern is regarding the existence of two statements within the question, those being "[. . .] you shouldn't think about your own experiences [. . .]" and "[. . .] you should just focus on what the book says." The entirety of the students who took a neutral position on this question (agreed and disagreed) centered their reasoning on these two statements. Consider the following:

> "I both agree and disagree. I think it is important to combine what you're learning from the book with your own experience to maybe learn more about what you experienced."

The example above focuses both on the value of personal experiences and on the value of knowledge within the text. As mentioned, this is a common theme among all who were neutral to this question. These students will disagree with "[. . .] you shouldn't think about your own experiences [. . .]" but then at least partially agree with "[. . .] you should just focus on what the book says." The emphasis on balance by the students here is not necessarily due to a poorly worded question, but simply student interpretation. That is, students do not want to dismiss the value of the text, therefore they do not wish to disagree with the idea that "[. . .] you should just focus on what the book says" despite their disagreement with "[. . .] you shouldn't think about your own experiences [. . .]." A possible correction needed to avoid this issue with students, then, is to restructure the question such that the two statements must be simultaneously agreed or disagreed upon. In other words, the question must accommodate students who both value the book and value making personal connections. As such, this may be a question which is best suited for part 3 of the EBAPS, where two students are making arguments for and against a belief. Since the question posed is similar to EBAPS question 1, we may expect that it is intended to mostly probe "the nature of knowing and learning". However, beliefs about science can be influential in student responses:

> "I agree with this statement. I believe that science isn't a biased process. Science is based off of facts, experiments to test hypothesis, studies, research and asking questions. To learn science you have to understand it the process and how tests or facts influence outcomes."

This student has a relatively sophisticated view regarding how information within science comes about and thus gives heavy credibility to the knowledge that science puts forth (thus, deep appreciation for the knowledge presented within the text). It is interesting, then, how these sophisticated views regarding science lead to an unsophisticated (agree) response at it pertains to this question. More accurately, the student does not appear to be weighing the importance of making personal connections to the material but is rather focusing directly on the credibility of personal experience versus the credibility of scientific evidence. Much like the previous concern, the question also appears to be influenced by beliefs about the real-life applicability of science outside the classroom:

> "I disagree because your experiences are based out of the real world and they are based off of what is actually happening."

Here the student is attributing their response to their beliefs about the real-life applicability of science (axis three of the EBAPS). The argument of the student is based around the premise that if science applies to everyday life, then of course personal experience will be linked to textbook knowledge by some means and should thus be valued. What is unclear is how the student would intend to utilize their own experiences alongside what is presented in the text. It is because of this lack of clarity that it cannot be stated whether this question is testing this student's view of "real-life applicability", "the nature of knowing and learning", or aspects of both. Further complicating this matter is whether it is the real-life applicability of science being tested or if that is instead a reflection of their beliefs about science (and/or, the nature of science):

> "I disagree with the viewpoint that a science textbook is the only right answers because science is made up of observations and experiences. So, it is only natural that personal experience comes into play."

As can be gleaned from another student statement above, views about the nature of science (based on evidence, i.e. observations and experiences) do seem to be motivating the student's value of personal experience. In this instance, at least, their views about science are interwoven with (or potentially creating) views pertaining to the real-life applicability of science, which then seems to dictate their views about "the nature of knowing and learning" as measured by this question. The last concern regarding this question includes yet another belief which may be capable of creating a noticeable outside influence, "evolving knowledge". One could postulate this affect well before seeing any student responses, given that student views about the tentativeness of scientific knowledge and their views about the content of knowledge within a scientific text are certainly going to be related. "Evolving knowledge" as presented here is dissimilar from one of the previously outlined concerns regarding beliefs about science and pertains to a student's view about the stability of knowledge (set in stone vs. completely ambiguous). To be explicit, beliefs about science above pertained more so to student views about the quality of scientific information as gleaned by the process with which it is acquired and not the stability of the information. As an example, a student may understand scientific knowledge as being our best effort towards accurate information, yet still believe that is capable of changing on a regular basis. So, finally, consider the statement below:

> "I disagree with this statement because science is always changing and one's own experience is due to the effects of science."

The latter part of the student's statement pertains to a previous concern (real-life applicability of science), but the former part of this student statement seems to be portraying a belief about the stability of scientific knowledge (that it is "always changing" and hence unstable). Thus, the heart of this final concern is how beliefs about the stability of scientific knowledge ("evolving knowledge") influence the question. The above student may have a rather unsophisticated view about knowledge in science (that it is highly unstable and possibly ambiguous) which would then lead them to disagree with the idea that "[...] you should just focus on what the book says." As has previously been the issue, it is not clear about the extent to which these particular views regarding "evolving knowledge" may be influencing their response. To give conjecture, it appears as though, for the student above, that an unsophisticated view about "evolving knowledge" and a sophisticated view about "applicability of science" gave way to a sophisticated view (disagreement) about "the nature of knowing and learning."

There were 47 total written responses to this question in all. Of the 47 total responses, 3 agreed (unsophisticated), 36 disagreed (sophisticated), and 8 were neutral (agreed and disagreed). The most prominent responses included the belief that personal experiences could be used as a study/learning strategy:

> "I disagree with this viewpoint because you learn best when you can make connections between the scientific information and your own knowledge and experiences."

> "I disagree with that statement. Our individual experiences can help us in learning science even if science disagrees with our current views or knowledge. If you look at anything with an open mind, you can learn. It's similar to meeting new people at school who disagree with you. You can still become friends with them even if they disagree with you or your viewpoints. You can also learn from them and their knowledge, thus building up your own knowledge."

These responses valued the role that prior experience can play when attempting to learn new material. Specifically, in how differences between personal experiences and textbook content can lead to the creation of new knowledge. These students did not focus on which information was more correct, but merely on how to utilize potential contradictions in knowledge to generate new knowledge. Another notable bit of content within responses was the belief that anything may be questioned:

> “DISAGREE: science was developed on questioning EVERYTHING”

> “I think that you should focus both on the book and your own experiences. You should question the book”

> “I disagree with that. I believe looking into your own experiences can help you question or understand what the text book has to say. You should not, not question something you don’t understand, because you can try and solve your question and get a deeper understanding.”

Students here, as outlined in the above examples, associate a strong value towards questioning knowledge. The nature of where this desire to question comes from can only be speculated. Within the first example the student bases the need to question on their views about the scientific process (it would seem), whereas the third example appears to treat questioning as more of a learning strategy. The next pattern seen within some student responses was that of attitude towards learning:

> “I agree to some extent to this question. Our personal beliefs should not make us biased towards science and if the book is showing reasonable evidences, then we need to put our personal beliefs away and trust science”

> “Yes, I do agree. I feel like there are certain topics that may not always a line with science like a religion. It’s important to see both sides but be able to understand science also”

> “I believe that science is a subject that we cannot just be spoon fed. It takes exploration and critical thinking to get past any biases that someone may have. So while our beliefs shape our lives, we also must keep an open mind to other possibilities, even if we don’t believe them.”

The trend here is that these students are accepting of at least entertaining other knowledge which may conflict with their own. More specifically, these students are concerned about trying to keep an open mind when learning. Although this trait may be present amongst many other students, individuals with this code explicitly exhibit open-mindedness within at least part of their response. Another aspect of student responses that will be focused on was also discussed earlier as a concern, that personal experiences are reflections of science

> “I disagree because learning science should help you better understand your own life and life experiences.”

> “I disagree because I think that science can be easily applied to our everyday experiences.”

The belief that personal experiences are important when learning science because science attempts to explain these experiences is a key focus for these individuals. This code has been already discussed extensively in the above section regarding concerns about this question. One final part of student responses worth discussing was the belief that some form of further reflection and/or analysis should be taken when encountering conflicting information with personal beliefs. In general, students either simply claimed that they should personally do more research:

> "I think that you should try to use your experiences to try and explain what the textbook says. Do some research and see what the facts say."

seek advice from a knowledgeable source:

> "One does need to focus on what the book says, that is true. But they should be able to take their personal experiences into account. And maybe, not having mastery of the subject, the two do correlate but the student has no way of currently understanding this. The student should consult the professor (This makes it important to have good professors)."

and/or refer to multiple sources (for confirmation):

> "I disagree. Anecdotes should not be taken in place of studies' results, but neither should a single source of information be blindly followed."

The overarching theme present within these responses may be best described as the nature of learning. Specifically, the value that students place on relating personal experiences to material. The emphasis is placed on learning and not knowing because students appear to view making personal connections as a means by which to acquire new knowledge, to learn. This is compared to the nature of knowing, which involves more metacognitive activities centered around student ability to actively monitor and reflect upon their understanding. Furthermore, the context of this question implies to the students that they have already identified a conflict in knowledge.

EBAPS Question 3

The question as posed to students was:

> "Obviously, computer simulations can predict the behavior of physical objects like comets. But simulations can also help scientists estimate things involving the behavior of people, such as how many people will buy new television sets next year. Do you agree or disagree with this viewpoint? Explain your thoughts."

This question is effectively the same as EBAPS question 3.

There are several concerns regarding the manner in which students are responding to this question. For example, some students focus on the example "[...] such as how many people will buy new television sets next year" and weighed their response heavily off of this:

> "I agree cause marketing techniques have used similar simulations to predict things for years so it makes sense that its possible."

Another mild concern regarding student responses to this question is that some also do not seem to be appreciating the term "estimate" and instead focus on the simulation getting exact results:

> "I disagree, you can take the sales report of how many new televisions were sold last year. Then estimate how many will be sold for this year. I don't believe you can get an exact answer."

As this question, like its EBAPS counterpart, should be exploring the applicability of science to real life, the detraction from real-life applicability due to the student interpreting an extreme case is unfortunate:

> "I do agree that simulations can guess the behavior of humans but since the human brain is so complex and complication these are merely estimations and not exact predictions"

> "I do not agree with this, I think people are far too unpredictable to have their choices mapped out by a computer, it seems unrealistic."

Overall, it is also possible that the question should move away from a tense which implies whether or not scientists can do these kinds of simulations and put more focus on gauging the extent to which science could assist in these types of estimates. Perhaps, a question more like "Scientists can estimate the behavior of things such as comets using computer simulations. Scientists can also be very helpful in estimating the behavior of different types of systems, such as the behavior of humans." This example question removes the use of a specific example within it (which students often fixate on, creating very unique opinions and/or general complications) and also attempts to put more emphasis on the importance of the applicability of scientists outside their traditional realm. Furthermore, this example question strives to take some focus away from what computers are capable of as compared to what scientists are capable of.

Of the 81 total written student responses 59 agreed, 13 were neutral (agree and disagree), and 9 disagreed. The most frequently coded aspect of student responses was what one may have expected, that simulation results are estimations of the actual:

> "I agree because scientists can study the trends that humans tend to follow. Sociology is considered a real science and is simply the study of humans. Also, these simulations claim to be estimations rather than actual predictions."

> "I agree. Simulations can help people see a hypothetical situation, although not always 100% accurate, they will mostly be accurate."

During the coding process, the lack of focus on this concept elsewhere was found to be a curiosity worthy of noting and hence why it was made a code. Another frequently seen aspect to student responses was the belief that people are complex:

> "I do agree that simulations can guess the behavior of humans but since the human brain is so complex and complicated these are merely estimations and not exact predictions."

> "I do not agree with this, I think people are far too unpredictable to have their choices mapped out by a computer, it seems unrealistic."

Out of those who disagreed, 7 of the 9 mentioned the unpredictability of students within their responses. These students who are disagreeing seem to focus on the accuracy (or lack thereof, rather) of a simulation and, occasionally, the capabilities of a computer. Instead, we would like these students to have focused on if scientists could be helpful outside their traditional domain, the former of the examples immediately above better reflects this view while yet stating their concern with human complexity. Another code was associated with a focus on computers, namely their capability:

> "I disagree with that. I think people are dynamic and they are always changing and it is ignorant to think a computer can always predict how people will act."

> "I agree. Computers are really good at predicting things and even better at statistics."

> "I agree with this. Computers can tell us lots. They can estimate based on research, how many people will buy new television sets next year. There is nothing wrong with this. It is not harming others, it is simply making a prediction."

Whether they agree or disagree, responses such as these student examples detract from the focus of the question. What is desired here is that students focus on the capabilities of a scientist, not their physical toolset. This is important, as 12 of the 59 who agreed did so with the capabilities of a computer in mind. As such, this should also show up as a concern regarding the question as a whole. Another prominent component to many student responses was a focus on the importance of past observational evidence:

> "Simulations have to be based on some form of prior research in order to correctly understand the fluidity of a market or idea. Understanding previous behaviors provides more insight into the particular situation."

> "I do agree somewhat with this statement because scientists can see how many people will buy a tv set because they can use numbers from last year to project those sales."

These responses are potentially shedding insight into student views not only about the usefulness of scientists, but how they can be useful. The last portion of some student responses was the reliance on outside examples or previous experience regarding the applicability of scientists outside their domain:

> "I agree because people are predictable and that is what psychologists study sometimes."

> "I agree. I've seen videos of sheep passing through a gate from above and they looked exactly like gas physics. Also, human behavior is predictable, like animal migrations or feeding habits."

> "I agree with this viewpoint. People act just like other objects, and are typically predictable. People are so predictable that the estimation of how many people will do some specific thing this year will be pretty close. This estimation process is even used by McDonalds (McDonaldization) to estimate the number of customers in order to have the right number of employees working and how much production will occur."

Although one might prefer that students realize this on their own, without personal prior experience or knowledge, at least they acknowledge the potential for this type of simulation to be possible. There could be concern regarding the extent to which these students are connecting the plausibility of the proposed experiment to the applicability of scientists outside their traditional domain, however. This concern regarding the use of an example within this question statement has been voiced above and is, at the very least, something to be aware of.

The predominant theme of this question seems to be student views regarding philosophical aspects of science. Views like science as being based on observation, science as having limitations, and science as attempting to be precise. From a broad overview, one may say this question tests the real-life applicability of science, but look a bit deeper and prominent nature of science views become apparent.

EBAPS Question 4

The question as posed to students was:

> "Discuss the extent to which you agree and/or disagree with the statement below. Be sure to explain your reasoning. When it comes to science, most students either learn things quickly, or not at all."

This question is designed to motivate similar responses to that of EBAPS question 4.

There is fundamentally only one concern regarding the manner in which students are answering this question. The question, and likely its EBAPS counterpart, does not claim to be associated with any particular EBAPS axis. The complication occurs, then, in that students associate learning things quickly with the word "natural ability" and hence this question becomes similar to the axis "source of ability to learn", as it probes an aspect of mindset. Student interpretations of natural ability are not all alike, unfortunately. Students have been observed to associate natural ability (learning quickly) with things such as past experience and/or interest, instead of with an innate born ability. Those associations have been discussed previoulsy, but will also be discussed below as it relates to this question. This inconsistency in the origins of natural ability can create fluctuations in how these students respond to the question within the current intended epistemological framework.

In all, 185 students gave written responses to this question. Of the 185 total responses, 10 were neutral (agreed and disagreed), 15 agreed (unsophisticated), and 160 disagreed (sophisticated). The most prominent belief seen within students was that some people learn faster than others:

> "I disagree. Students all learn at different paces. Some may get it immediately and other may take longer but no one is completely unable to learn science at all."

> "I disagree. I think that when it comes to learning anything all that matters is how much you are exposed to it. Some people only need to hear something once in order to learn it. While other students may need a few extra reps to understand the material, but it is possible to learn it even if you don't learn it as fast as most other people."

> "I do not agree with the entirety of this statement. I do agree that some kids learn science much more quickly than others. It comes more natural to them if they have been exposed to a great deal of scientific minds. However, I do not agree with the not all part. If someone claims that they know nothing about science then they are purely just lazy. While many concepts in science can be very hard to wrap your head around, basic processes and ideas are fairly easy. There are many things that do not take a great deal of effort to learn."

> "I think it depends. I mean if you have the interest you could learn things fast because you care and it matters to you, you enjoy it so it is easier to learn faster because of incentive drive."

In the responses of students within this code, most were similar to the first example above in that they did not specify why they believed that different people learn at different paces. Of those who did specify, there was a variety of responses which seem to be focused around the different definitions of "natural ability" that have been presented elsewhere. The second example attributes the different learning speeds to an individual's learning style. Whereas the third example discusses prior experience within science and the fourth example links learning speed with personal motivators (both are subsets of "natural ability" definitions). Few were the cases where learning speed was explicitly connecting to intrinsic natural ability:

> "I completely disagree with this statement because every student has a different way of learning, which in the end, contributes to how fast or how slow they learn the material. Some students in a classroom already sort of have a grasp around the material or they are just naturally fast learners so the material comes very easy and quickly to them [...]."

The next prominent component seen in student responses was that of the utilization of a personal account:

> "I think this is true especially for my experiences in school. For me science does not come easy and for a lot of my friends, especially in high school it was very easy for them."

> "Based off of personal experience I completely agree with this. Through out my schooling when it comes to science there are concepts that I understood right away and there are concepts that I never understood. It was extremely rare if there was ever something that was in the middle of those two extremes."

> "I don't agree with this statement because I am one of the people who didn't learn science very quickly but by the time the test came around I understood what was going on. I think science is just like many other subjects where you have to work hard to understand the concepts."

> "I would disagree with the statement because all things take time. If you aren't learning something quickly, then you go to a professor and ask questions, get a tutor, and practice the concepts with homework."

In students who agree with the statement (first two examples above) you usually see an emphasis on the difficulty of the material or a comparison to how others are having an easier time with the material. This is contrasted by students who disagree (latter two examples above), where they usually discuss the initial difficulty but then explain a means by which to overcome that difficulty. From an epistemological standpoint, students could benefit from groupwork by seeing their peers also struggle with the material. It would be interesting to compare and contrast these individuals study/learning habits.

The remaining four factors seen within student responses all fall under a more general category involving the idea that other variables, beyond natural ability, are influential in learning. The first to be discussed is that of hard work:

> "I disagree, there are many scientific things that have many levels such as calculus. They take time to learn and understand."

> "I disagree with the statement. Some students learn faster than others, but eventually most people who put in the effort to learn will understand science."

> "False. Learning is something that you have to practice - we don't live in the matrix, and while we do have a super computer within our heads, it takes a lot longer then a 30 second YouTube video to understand the vast complexities of the universe."

> "That is not true, anyone can learn science concepts it may just take more work for others. This question is similar to the question, in the beginning, asking what is better someone who is naturally gifted or a hard worker. The information may come easier to the naturally gifted, but the hard worker will come to understand what is being said because they will work for it and not quit."

Aspects of these responses tend to focus around hard work, that being effort and/or time put forth in an attempt to understand the material. These individuals believe hard work to play a key role in learning. As seen in the third example above, practice (and studying) were also grouped with this code. This was done because the manner in which students utilized these words appear to embody the idea of an effort and/or time commitment, as compared to an actual method within their practicing and/or studying. The second of these factors is the importance of maintaining hard work:

> "[...] I don't think that anybody can't learn anything. We just have to work harder, and have lots of determination. That makes the end all the much sweeter."

> "Neither. Science is hard to grasp. It takes work and dedication to master concepts."

> "[...] So I believe that learning is all about how much work a person is willing to put in and as long as some people are ok to not understand it right away and to be patient and persistent then they can understand the same things that a person who learns fast can [...]."

Focus is being placed on commitment here, with students using key words such as "determination", "dedication", and "persistence". As can be seen in the examples, these traits typically complement hard work. A surprising, yet sensible, third factor that students were seen to focus on was if they cared:

> "Students learn what they feel is most interesting to them. There's some things I remember from my elementary science class and there's somethings that I don't."
>
> "So I definitely agree with this statement and science is one of those things that you either just think its cool but don't care about the in depth explanations of it or you love it and need to understand why things are the way they are. I think that is the main difference between students learning things quickly or not at all."
>
> "Some people may be more adept to learn quickly, but with enough work and commitment and passion, anyone can learn it eventually."

These responses had students expressing the importance of emotion. Specifically, students believed that feelings such as passion, interest, and curiosity play a crucial role in an individual's ability to learn the material. It is unclear within responses whether students associated these factors with increased motivation, determination, or hard work. In general we can likely say that these students believe that if someone has an interest in science that they are going to put more effort in and thus learn the material more quickly. Many paths leading to what it takes to succeed do seem to include hard work, regardless of the role that commitment, interest, and/or motivation may play in getting there. The last of the four factors is associated with learning methodologies, be them pre-class, post-class, or during class ("class" could, in a more all-encompassing sense, be exchanged with "exposure to content"):

> "I don't believe that is true. If an average someone doesn't understand a scientific concept (which is appropriate for his level), then he simply didn't find the right learning style or explanation of that concept."
>
> "I disagree with the statement because, while some students take very quickly to science, the ones who struggle at first can often be taught a new way to look a problem or concept and then understand the subject after doing some practice and studying."
>
> "I disagree but at the same time I partially agree with this statement because in science, being the complicated subject it is and the density and depth it goes or that we can take with it a certain depth, it mostly depends on the teacher you have teaching the material to you that will determine if you retain the information."

The common trait within this factor is the reflection on learning and why one may be struggling to grasp the subject matter. Fundamentally, all these responses hinge on the belief that learning occurs differently for different individuals. This ties in nicely regarding potential reasoning as to why certain people learn the material more quickly than do others.

The overall theme regarding student written responses to this question was centered heavily on the role that natural ability has with student learning. It is likely due to the observation (which has been previously made) that many students define "natural ability" as the ability to learn things quickly. That is, students may read this as "when it comes to science, if you do not have natural ability then you will not learn the material." Because these particular students responded to a hard work versus natural ability question before the current question, some students even acknowledged a similarity between the questions, with one student saying: "This is similar to the work ethic vs natural ability question." This question is, however, different from those other questions on the EBAPS (questions 21 and 22, for example, as gauged by written student responses) which seem to directly oppose natural ability and hard work. Within this question natural ability and its influence on what can be achieved is at the forefront of student thought, with hard work being used a justification for an alternative learning route. Due to the focus on natural ability, more insight is gleaned into factors that may be responsible for learning the material quickly, such as personal interest in the material.

EBAPS Question 5

The question as posed to the students was as follows:

> "If someone is struggling in a physics or astronomy class, do you believe that studying in a better way can make a big difference? Explain your answer."

The only difference between the EBAPS question five and the above question is the exchange of the word "chemistry" for "astronomy".

There are two primary concerns regarding the nature of this question. The first concern pertains to the wording of the question itself, specifically the phrase "studying in a better way can make a big difference." The use of the word "better" within the previous phrase implies kind of circular logic that students can detect:

> "[...] Studying in a 'better' way by definition would be better [...]."

Hence, students who identify this are more likely than not to follow through with a sophisticated response. Responses from students do indicate that the majority of students are associating "better" with "different", which is an acceptable substitution. Very few student responses highlighted the second concern of this question, that being the context of the question. This concern is best highlighted with the following:

> "It depends on why they are struggling. If they just don't understand the concept then they should have someone further explain in a way they would understand, but if they are struggling with tests then they should change the way you study."

As you may see, the student response is dictated by whether one is actively attempting to make sense of the material, as if they are first trying to learn the material, or whether one is preparing for a test of their knowledge on the subject matter. Most students seem to have keyed in on the former portion of question five "if someone is struggling in a physics or astronomy class [...]", taking "struggling" to imply an attempt at actively trying to learn the material (as opposed to attempting to re-familiarize themselves with what they already know). Epistemologically, the researcher must be aware of the possibility that beliefs associated with whether studying in a different way is beneficial may differ depending on if the students are first attempting to understand/learn as opposed to if they are reviewing material that they already know. In future epistemic work, researchers should also take care in their use of the word "study" or derivatives of such. It is possible that in this question the word "study" was associated most strongly with an attempt at actively trying to understand the material due to, we propose, the use of the verb "struggling" also present within the statement. Without that context provided, this question may have brought forth more responses geared towards behavior when preparing for exams instead of beliefs associated with their ability to learn.

The majority of the responses to this question are sophisticated (in agreement to statement), with only six of the 81 total coded responses being unsophisticated. Given the student interpretation/s of the question, the overarching view would be: students agree that if you are struggling to understand the material, then attempting to learn material in a different way can be beneficial. The most frequently coded responses were in line with this view, focusing on either the approach to a new learning methodology:

> "I think so because if he or she is just looking at the lecture notes, it could be more helpful to read the book or textbook. It opens up a new way of getting the same information [...]."

or put emphasis on the effect of an alternative learning methodology:

> "I think that changing the ways in which you study can really help students, because it changes the way you learn and think."

Of the responses which were coded based on change in learning methodology, the most prominent method outlined by these students involved social learning. By "social learning", we refer to either a mention of studying within a group or seeking other students so as to have it explained in a different way:

> "Yes, studying [...] in groups can increase understanding and retention of a subject."

> "You need to go get help from someone who knows what they're talking about to help you. It is much better than trying to struggle through material by oneself."

Also prominent here was an explicit mention of asking for help from a source of authority such as a tutor, teaching assistant, or the instructor:

> "Yes, I believe changing study habits can help, but I also think that getting help by a professor or TA would be the best course of action."

Numerous alternative strategies were also mentioned such as note-taking, reading the text, and working problems:

> "I think so because if she or she is just looking at the lecture notes, it could be more helpful to read the book or textbook. It opens up a new way of getting the same information and could somehow be clearer."

The responses which emphasized the effect of alternative learning methods lacked the clarity seen above, with many students simply saying it would "improve understanding":

> “Yes I believe that studying in a better way can make a difference. The method of studying can greatly improve their understanding.”

What is meant by “understanding” and what it means to “understand” something was not present. Of the trends which were present was one in which changing learning method would make understanding simpler by improving accessibility (number of available learning methods) or perspective (attempting the learning approach of another individual):

> “Yes. Everyone learns in different ways. By branching out from the typical methods, one might be able to find a better study method more suitable for their mind.”

> “I think that changing the ways in which you study can really help students, because it changes the way you learn and think.”

Another sub-code within the effect of alternative learning methods was that of retention and familiarity. Several responses mention that different methods could help in that regard:

> “Yes, the way that you study can directly affect the way that you retain information.”

The two prominent codes discussed above, fall under the general attitude that learning method/approach can make a difference (learning/understanding). There were several other codes that showed up frequently enough to be worthy of mention. The first was the “hard work” code. Here, responses coded for hard work would usually directly mention things such as “hard work”, “effort”, or “time” within their response:

> “[...] everyone can learn it if they put in the effort.”

> “Depending on the person however, learning physics or astronomy can take hugely different effort levels [...] but eventually given ample time I think most human’s can learn these fields.”

The above coding was made distinct from the coding associated with learning methodology, as precisely what is meant by things such as putting in “hard work”, “effort”, and “time” isn’t clearly put forth by the student. It may be more likely than not that these responses are also entangled with student views regarding learning/study methods. Yet another code which showed up with relative frequency emphasized the belief that everyone learns/studies differently:

> “Yes, studying works differently for all people, some forms of studying are better than others, and one should find the way which works best for them.”

> "Yes the way you study can make a big difference. People learn in many different ways so if you aren't understanding something switching the way you're studying it can help a lot."

Lastly, students responding to this EBAPS question would also be influenced by their personal experience, typically mentioning this such as:

> "No matter what class, studying is a sure way to improve. Studying in different and better ways definitely makes a difference, I know this from my own experience with studying for this class."

> "Yes, because for my first test I didn't study a whole lot and got a pretty bad grade. But for my second test I changed my study habits, and got a much better grade."

The dominant theme present within these responses was what we may call "study and learning strategies". The vast majority of students, when responding to this question, focus on some aspect of personal study/learning strategies and their efficacy, but did not consistently reflect upon any specific strategies.

EBAPS Question 6

The question as posed to students was:

> "When it comes to controversial topics such as which foods cause cancer, there's no way for scientists to evaluate which scientific studies are the best. Everything's up in the air! Do you agree or disagree with this viewpoint? Explain your thoughts."

This question is intended to be a reflection of EBAPS question 6.

The biggest concern regarding student responses to this question is the choice of example for controversial topic. The majority of students who respond in an unsophisticated manner (agree) with the above statement focus on the difficulties present in determining what causes cancer. Fundamentally, it would be ideal for students to focus on if it is possible to make distinctions between any conflicting scientific findings. How one might separate views regarding that distinction as it relates to a given topic, let alone a controversial topic, is unclear. It could be said that these students are answering as intended, that they are exhibiting an unsophisticated belief with respect to analyzing conflicting scientific information. However, it is not unreasonable to state that these beliefs depend critically on the topic in question. For example, these same students responding negatively to the above question may respond in a more sophisticated manner when topics are varied: "When it comes to controversial topics such as how much water one should drink a day [...]." The point being, perhaps the topic of cancer (or any topic, for that matter) reflects a unique view as opposed to a more general belief regarding the evaluation of typical (however one may wish to define "typical") scientific claim/s. To elucidate the degree of student sophistication (or lack thereof) regarding this view, it would be enlightening to ask this same question several times while varying the degree of topic complication. For practicality it may be best to frame future questions to be topic independent: "Independent groups of scientists put forth two different studies regarding a controversial scientific topic, yet both studies have different results. Is there any way for scientists to evaluate which study is the best?" A related curiosity regarding EBAPS question 6 is the use of the word "scientists" as opposed to "you", as one puts emphasis on the capabilities of scientists and the other would put an emphasis on the self.

Of the 49 total responses to this question, 35 answered sophisticatedly (disagreed), 12 agreed, and 2 were neutral (agree and disagree). Codes for student response were numerous and, hence, sparsely populated. As a result, code categories will be discussed, alongside the actual codes which populate the code theme. The most prevalent category is what we call "philosophy of science". Students responding here would focus on what science does and/or what it tries to do. Written responses in this category include views such as science attempts to be precise:

> "I disagree through testing and statistical analysis scientists can make conclusions with extremely high levels of confidence. Therefore some controversial topics can be decided one way or another."

science is based on evidence:

> “No, I don’t believe this viewpoint. There are a lot of things that scientists don’t know and really is kind of up in the air, but they do have good evidence for many things. There is strong evidence that can prove and support new theories and ideas.”

science progresses gradually:

> “I disagree. Even though the state of knowledge in some fields can change frequently and significantly, scientists continue to push us closer to a correct understanding by evaluating and accepting/rejecting research which has been done.”

and the process of science:

> “I think that as long as the studies are conducted well, they should be taken into account.”

> “I disagree. This is the type of question that can be answered via controlled studies or historical data collection. Even if a conclusive answer can’t be found now, the possibilities can be narrowed down.”

Responses involving “science attempts to be precise” take into account the idea that science is doing as good a job as can be expected to yield the most accurate results. Responses with “science is based on evidence” focus on the role and importance that evidence plays within science, presumably as distinct from opinion-based reasoning. “Science progresses gradually” is a view expressing how, overall, science continuously moves toward a more concrete understanding of a topic. Lastly, “process of science” responses allude to analyzing how the science was done. Unfortunately, few students discuss beliefs regarding how they think science should be done, revealing only that they put value on how the study was conducted. The example above discussing “controlled studies” is a rare instance of a student giving insight into their views about the process of science. The next most prominent coding category was that focusing on the requirements of scientists. This category has students discussing what they believe scientists depend upon in conducting their work. Two subcategories are present here, “peer review”:

> “I disagree. Of course scientists know which studies are best. They’ve done the research, the experiments/studies and have had their work reviewed by their peers who may or may not agree. They’ve done the work and presented their findings. It’s up to the public on how they handle the results.”

and “replicability”:

> "If a study can be reviewed and recreated with similar results by peers, these studies should be held to be more true."

"Peer review" codes refer to students who emphasize the importance of having a study be reviewed by other scientists capable of doing so, while "replicability" refers to the view that the finding/s of a study should be capable of being reproduced. The next category to be coded was that in which students seemed to focus on personal assessment techniques as opposed to methods scientists could use. This involved looking at the opposing argument and/or the reliability of the source:

> "I disagree. There are people who claim will present studies that are biased, and created to get one specific result. There are non biased ways to evaluate studies such as how the information in study was researched and aquired. Did they look at the opposing argument? Did they only use biased sources?"

> "I disagree with this viewpoint. There are lots of factors to look at to see if a study is credible from who paid for it and who's running everything to how they go about doing it."

As can be seen, students here appear to take the question to imply how they personally can establish which scientific study is the most reliable. Usually these responses involve wanting to know who did the study, who reported on the study, and who was responsible for the study. A code which arose primarily due to the students who agreed with the statement was that of problem complexity:

> "I agree because it is hard to tell if any foods do cause cancer and the scientist are doing their best to find evidence."

> "Finding if certain foods cause cancer would be a hard task. Not only do you need to find a place to start, but who would you test? Everyone has different diets, different tastes, etc. So keeping an experiment and then applying it to a society would be probably impossible."

These students put emphasis on the difficulty involved in the study, it is unlikely that the results can be taken with confidence. Note that not a single response which was coded for "difficulty" was done so without also focusing on the topic of cancer. However, students were seen to focus on the topic of cancer without explicitly focusing on the difficulty involved in obtaining reliable results:

> "I agree. Most of the time saying a specific something causes cancer is a shot in the dark. Don't get me wrong, there are some things that definetly cause cancer, but there are very few."

Perhaps it is more accurate not to say that problem difficulty wasn't previously addressed by those who disagreed, but rather, students who disagreed with the statement were dismissive of the difficulty involved, as opposed to those who agreed:

> "I disagree, scientists can look to the tests with the least amount of variables, and the strongest evidence. Some tests are very inaccurate because of outside variables."

This student is an example of a sophisticated response in which students acknowledges the complexity (difficulty) of a problem, but finds a means by which to overcome it. Students who claimed "difficulty" as a reason for statement agreement, typically also refer to the large number of variables present in the topic:

> "I agree because there are a lot of other factors that could cause the cancer, not just the foods themselves."

Of the students who agreed with the statement, 7 of the 12 focused on the difficulty involved due to the topic of cancer.

The prominent theme within the student written responses was the focus on their views about science. Views about how science and scientists operate. In general, the philosophical and sociological aspects of science and scientists. One could say that this question indicates a degree of sophistication regarding several tenets of the nature of science and/or the capabilities of scientists.

EBAPS Question 7

The question as posed to students was:

> "Discuss the extent to which you agree and/or disagree with the statement below. Be sure to explain your reasoning.
>
> A teacher once said, 'I don't really understand something until I teach it.' But actually, teaching doesn't help a teacher understand the material better; it just reminds her of how much she already knows."

This question is designed to motivate similar responses to that of EBAPS question 7.

The primary concern regarding the manner in which students responded to this question is the apparent dichotomy of the question itself. The two parts of the statement students focus on are "[...] teaching doesn't help a teacher understand the material better [...]" and " [...] it just reminds her of how much she already knows." Consider the following student examples:

> "I guess that teaching could also be a reminder to someone of how much they already know, but in my own experience, this is not the reason why teaching has helped me to better understand material. Teaching has helped me to better understand material because it has forced me to learn the material myself."

> "I don't agree with that statement. Though teaching does have a direct correlation with understanding it, understanding it does not always require teaching it. For example, I understand Dutch and German and English; I can speak the languages. But, I don't teach it. Teaching something to someone is beneficial in many ways, both to the teacher and learner. However, you can do one without the other; you can understand something but not HAVE to teach it to remind someone how much he or she knows. When I think of the term 'remind', I think of knowledge that has been forgotten, but then if told again, it 'rings a bell'."

The former example nicely portrays the student acknowledging the benefits of actively working through material and reflecting upon their understanding. However, they are mildly distracted by the part of the statement " [...] it just reminds her of how much she already knows." The latter example then shows a student who has gotten completely sidetracked by that same aspect of the statement that the previous student did. These two statements need not be diametrically opposed, as the previous student mentioned. Furthermore, students can become confused about how to treat " [...] it just reminds her of how much she already knows" when they agree with "[...] teaching doesn't help a teacher understand the material better [...]":

> "I agree with this statement but I feel it is misleading in how it is worded. While teaching something to others does remind you of what

> you know it also forces you to put a concept into words and in a way that someone unfamiliar with the topic can understand. I completely feel that I understand a topic better by trying to explain it to someone else. Generally if I can't put what I am thinking into words then that means that I don't have a great understanding of the topic myself."

As can be seen, the student wishes to partially agree with the latter part of the question but acknowledges the conflict that agreement creates with the former part of the question. In essence, the tone of the question is conveyed as there either needing to be "understanding which occurs" or "recollection which occurs". This has been seen in other written response questions and there may need to be a deeper consideration about the placement of two extreme statements within one overall question. Whether these observations are of any deep concern is unknown, but it is certainly worth mentioning. Nevertheless, many students appear to answer in a manner which indicates that they are considering the value of working through material and reflecting on previous knowledge.

Of the 190 total written responses to the question, 18 were neutral (agree and disagree), 47 agreed (unsophisticated), 117 agreed (sophisticated), and 8 were unclassified due to a lack of clarity within the responses. The first trait seen within student responses was the belief that teaching requires mastery prior to class:

> "I agree with the teacher's statement. Until you can teach something you are not a master at the material. Once you can teach someone what it is that you know then you become a master over the content you are studying."

> "I agree with the statement. Teachers must fully understand something before they teach it. Teaching may help develop an even deeper understanding, but one must completely understand before being able to teach it."

> "I think to properly teach students; the teacher must have a mastery of a subject. You can clearly see when the teachers know how to teach well because their students have good exam scores. I think that teaching in multiple ways would a good example of the of the quote above."

It may not go as a surprise to know that 18 out of the 47 agree responses contained at least a portion of their written statement exemplifying this belief. It was also seen within the sophisticated answer, however, yielding a single response of this type:

> "I disagree completely. To teach something correctly, you have to understand it fully, and you have to go through the steps of the understanding when you teach, which gives you familiarity with each piece of the material."

As can be seen, this response fundamentally differs from those seen within the unsophisticated responses. Within this response, the student still believes that the instructor needs to have mastery of the subject, but that in preparing for the class they obtain a deeper understanding and hence become proficient. Another code involved the idea among students that teaching strengthens knowledge that is known:

> "Teaching helps solidify something she knows even more. You can know certain things but will eventually forget. By doing something over and over it helps to remember exactly the answer that your looking for."

> "I agree not with the quote but with the statement. If a teacher doesn't have an understanding of the topic he or she is teaching, there is no way that it will be clear to the students. Teaching affirms and strengthens your already existing understanding of something but it doesn't give you understanding."

> "I strongly disagree with this statement. When we relate information about a concept to someone else through writing or verbally, we force our brain to make connections that may not have previously been there. By doing this, we solidify the information in our minds and expand upon what was already there."

> "Teaching a topic definitely causes the teacher to go more in depth with their research to expand their knowledge."

The key here is that these students appear to believe that teaching either strengthens knowledge that is known or it makes slight increases to their knowledge, like refinement. This aspect of students' responses became a code because of how they seemed to be viewing knowledge, saying things such as "adds depth", "solidify", and/or "expand". The intention was to separate these types of refinement/strengthening viewpoints from other viewpoints which acknowledged teaching as playing a more prominent role in improving understanding. Implication of improved familiarity or small additions to fundamental knowledge is not always clear in these responses. Some students did acknowledge that one may not be learning new material but can still "deepen" knowledge:

> "[...] In talking through it and having to answer questions about the subject, she'll comprehend all of the information more. While she may not be learning new things, she will be gaining a deeper understanding of the matter through teaching it and answering questions to others."

It is for precisely this perspective that this code was created. It would take further interviewing of the students to understand what their views are regarding "deepening" knowledge. The tone given by the student above would imply that "deeper understanding" refers more to making small additions to a set of knowledge that is already held. Similar to the previous code, but seemingly more sophisticated, is the student view that gaps in content knowledge can be identified and/or addressed by teaching:

> "I guarantee that no one is truly a master of anything. As one might be an expert, there is always room to grow. There have been many professors that are questioned, and they do not know the answer. They teach themselves, to better express the material in a better way. We as a human race expand our minds every day, that's why we are in college, isn't it?"

> "I think that that is an interesting statement. I think that teaching definitely helps people to understand topics better, because it forces them to fill in the gaps in their understanding and string together ideas in a cohesive manner so that they can explain the topic to someone else."

> "I disagree with the statement. It is hard to be a complete expert at everything in a subject. As the teacher teaches it refreshes them on what they already know as well as teaches them things that they had not learned prior to teaching."

The examples above indicate a similarity to the idea of strengthening/refining knowledge in that students view the instructors as already having a large amount of knowledge on the topic. The difference between the codes lies in the open acknowledgment by the students that there are now notable gaps in knowledge which are being filled/connected. Another aspect of some student responses is one which might be said to be an expansion of the previous, that teachers can learn from students:

> "[...]We are only able to teach what we know or what we know of. You cannot successfully teach someone calculus if you yourself cannot do it. Student questions can reveal to the teacher the existence of holes in their expertise, however."

> "I disagree, the best way to learn something is to teach it. A teacher can teach the same class and still learn something new everyday from their selves or the students even. It would be really hard not to understand the subject matter better while teaching it every day."

> "The statement above is understandable, but not concrete. I think that when you're teaching material you are often asked questions that further your own understanding of the material. I also think that sometimes saying things out loud can make you see them much clearer."

There still exists an envelope of instructor learning within these responses, but these students attribute the students themselves as being a productive part of the teaching, enabling the instructor to identify and reconcile gaps within his/her knowledge. The last trait seen within student responses that shall be discussed is the belief that new instructional strategies/explanations can be identified and possibly addressed:

> "I kind of agree with the statement. If you are able to teach something it shows that you really know what you are talking about, because you can articulate it in a way that other people will understand it too. On the other hand you could understand something perfectly but couldn't be able to teach it others because your way of explaining would be so different from how other people would understand it."

> "I think that teaching something could definitely help you understand it better. To teach you have to be able to explain something in a way that makes others understand that concept, therefore it should help you understand it better yourself. If you teach others you could say that you are teaching yourself as well."

> "I disagree that teaching material doesn't allow for someone to understand the material better. When you teach, you develop new ways of understanding the material and have the ability to think about the material in a new way, this can allow you to learn new things about the subject that you may not have otherwise thought about. Questions from students can allow the teacher to further their knowledge on the subject."

In the most encompassing statement, these views seem to frequently associate the teacher having to adopt new perspectives as them also gaining both increased content knowledge and furthering their capabilities in conveying that knowledge. It is possible that students are reflecting on their own learning strategies here and acknowledging that teaching can help further one's awareness of alternative learning strategies.

The overarching theme in these written student responses would seemingly be associated with the "nature of knowing and learning" from the EBAPS, as intended. Due to the presence of the student taking on the role of teacher, it is difficult to disentangle views about how knowledge may be created from views about the role of authority. Those with unsophisticated views appear to treat authority as holding essentially all knowledge possible on the subject. Those who answer sophisticatedly seemingly do a better job of placing themselves in the role of an authority. From here, these sophisticated students associate value with actively working through material both independently and with others. The value that is being referred to here is the ability to identify and expand upon gaps in their knowledge, as well as being able to make connections that had gone otherwise unnoticed. As such there may also be aspects of "structure of knowledge" from the EBAPS within these responses as well.

<u>EBAPS Question 8</u>

The question posed to students was:

> "Scientists should spend almost all their time gathering information. Worrying about theories can't really help us understand anything. Do you agree or disagree with this viewpoint? Explain your thoughts."

This question is intended to be similar to EBAPS question 8.

There exist several concerns regarding how students responded to the above question. By association to its EBAPS counterpart, the question should be probing the "structure of scientific knowledge". Written response data instead hints at the possibility that this question is primarily bringing forth students' ideas about science (theories, in particular) as well as the "source of ability to learn" (malleable mindset beliefs, specifically) and "nature of knowing and learning" (learning through engaging and reflecting upon material). It is suspected that this is due to the duality of the question. To elaborate further, this question makes two statements which appear to be related to the beliefs previously mentioned. The statement "scientists should spend all their time gathering information" seems to bring forth views about the role of scientists or, more accurately, scientific theories. The subsequent statement of "worrying about theories cant really help us understand anything" also puts focus on theories, but in the context of what "worrying" may accomplish ("nature of knowing and learning") and if "understanding" could be achieved ("source of ability to learn").

Of the 79 total written student responses, 56 disagreed with the statement, 9 agreed, and 14 were neutral (agreed and disagreed). There were two views/beliefs (codes) amongst the student responses which occurred most frequently. The first of which was the belief that theories drive or create a search for new knowledge/information:

> "I believe that scientists should spend a lot of their time gathering information because that is where evidence comes from. Though I disagree with the fact that theories cannot help us because they can. There is a lot of information already gathered within theories that can help further the research and information for scientists to study."

> "There is an important relationship between the use of information and data in the reasoning of theories and working on new discoveries. It is the contemplation of theories that leads to the specific collection of information and data."

> "I disagree with this statement. Theories are important to science because they challenge those who formulate them and others to explore new concepts and ideas. Theories are the basis for progress."

These students view theories as important, predominantly because exploring a theory can generate questions which function as a motivator for the pursuit of more information. The second most occurring view was the idea that "worrying" about theories builds upon our understanding:

> "I agree that science is a lot of gathering information and testing theories. I also believe thinking about theories can help someone better understand it."

> "I disagree with this viewpoint. Research can only get so much, and theory is then supplementary to the research to then extend the scope. Where research is lacking, theory helps."

> "I disagree completely disagree with that statement. The more important part of being a scientist is creating and developing theories to help us better understand the world around us."

Whereas the previous belief focuses on theories as guiding us towards new information, these student responses focus on theories as creating understanding. A facet of this coded belief occasionally arises where students will blend in a view about theories themselves:

> "I disagree with this viewpoint because theories are the basis of all knowledge. Without giving them the consideration and thought they need, no new knowledge/fact would ever be created."

This student response is clearly emphasizing that consideration of theories can generate new knowledge, but also states theories as the basis of all knowledge. So, it is not clear if fundamentally this student disagrees with the statement because of their views about theories, the value they place on interacting and reflecting upon theories, or both. This could be important because a focus on theory may be a manifestation of their views about science, whereas a focus on knowledge creation could imply views about mindset and/or the nature of knowing and learning. Another noteworthy view expressed by numerous students was that theories need evidence:

> "I would agree gathering information is how you solve a theory. If you don't do the research then you don't get the satisfaction of proving your theory correct."

> "Worrying about theories is important too, a scientist should consider all possible explanations to a phenomenon, instead of sticking to one and blindly refusing all others."

> "An idea used to account for a situation or justify a course of action seems to me to be a required element to a scientist's study, and the gathering of information is part and parcel of a scientist intending to explain something. It seems to me gathering information and 'worrying' about theories are exactly what is needed to help us understand things."

Students here have a part of their response which expresses the importance of obtaining information to test the theory. It is not surprising that 7 out of the 9 students who answered unsophisticatedly (agreed) to this statement included this view in their reasoning. Although the students who agreed to the question statement had responses which were founded on a sophisticated tenet of science, it was unsophisticated in how seemingly low a value they placed on "worrying" about theories. There is some division to the final code that shall be discussed. In a broad sense, some students hold a view that theories are generated from information:

> "I believe that scientists should spend most their time gathering information opposed to spending it on coming up with theories because I think spending more time on gathering information would give them a better understand on the theory they will be coming up with to make it more accurate."

> "I would disagree with this statement because I think that scientific theories are important and the point of gathering information is to help us form theories."

In some respects, this is like the opposite view that other students were seen to exhibit in which they believed theories to motivate the search for information. Instead, these students view information to be the motivator for theories. For now, this seems to be more of a curiosity than a concern, much like the argument of "which came first, the chicken or the egg?" A subset of this code had students expressing the same overall belief, but showing a bit more sophistication in their responses:

> "I disagree with this statement because the theories are in a way what put the info together, so everyone can somewhat understand the info given."

> "I disagree with that. All the information you gather is worthless if you don't look for tendencies and make theories about why something is the way it is."

> "I think that it is important to gather information, but I think it's just as important to analyze the data, and relate it to what you already know. Creating theories and trying understanding how things work is the reason scientists collect data."

These responses still fundamentally treat information as being at the origin of scientific study, as opposed to theory. Unlike the previous responses, however, these students are putting emphasis on connecting information to create a theory. By explicitly stating this, these individuals are displaying more depth with respect to their views about the nature scientific knowledge. So perhaps it does matter if the chicken or the egg came first, since students who believe information to be the progenitor of theory are seen to express more detail within their discussions of theory creation.

The theme of student responses for this question appear to be centered around scientific theory. Whether it be the belief that theories provide explanatory power, theories guide a search for knowledge, theories connect information, or even that theories are meant to be tested. Many of these views, with respect to the nature of science, would be acceptable representations of several philosophical aspects of science. Of course, a second (less prominent) theme does appear within the data, that being the consequences of concerning ourselves about theories. Namely, that reflecting upon and interacting with theories can lead to new or furthered understanding.

EBAPS Question 9

This question was posed to the students as follows:

> "If someone doesn't have a high natural ability in physics or astronomy, can that person still learn the material well? Explain your answer.

The structuring of the question is quite similar to that of its EBAPS question 9 counterpart. There is a discrepancy in that this question does not explicitly mention a class/classroom, but responses seem to have assumed this anyway. Furthermore, "astronomy" has again replaced "chemistry".

There are few concerns regarding the nature of this question. For the most part student responses indicated no confusion in what the question was asking. There does exists a concern that students who do not necessarily believe in natural ability and/or do not understand what is meant by natural ability may not be best expressing their personal beliefs in the manner intended. Below, a student acknowledges their confusion regarding the word "natural ability":

> "Yes, I'm still not sure what 'natural ability' for science means. I believe that any person capable of learning physics or astronomy can learn it if they try."

The entirety of the responses were sophisticated, that is, all 65 written responses were of the belief that someone can still learn the material well in physics and astronomy even if they do not have a high natural ability. The most frequently coded response tied hard work in with natural ability as students would make a general statement regarding how individuals without high natural ability can make up for it with hard work:

> "That person can absolutely still learn the material well. They just need to work harder than someone with the natural ability."

> "I think the person can still learn the material well, they just have to put in more effort."

Yet another response type which was as prevalent as the previous coding was the need for hard work:

> "Yes. You can always learn the material well it just takes time and practice and eventually you can learn the material regardless of how much time it takes."

> "Yes. If they put in the effort and work hard, they can understand the material."

This type of responses discusses hard work as purely something that is necessary to learn the material. These responses are distinct from the others above in that they do not clearly make any reference to the role natural ability plays, if any at all, beyond their agreement to the question. In the context of the question, it is possible that any number of these students are also expressing a belief that hard work can make up for not having high natural ability. To clarify the need for this distinction, agreement to a statement may mean agreement with a portion or version of the statement. In this case, students may simply be agreeing to the idea that anyone can learn the material well (regardless of any role that natural ability may play). A similar, but much less frequent focus, was that of resolve (commitment to learning):

> "Even if you aren't naturally gifted in a field you are studying within, you can still learn the material well. Hard work and determination can get someone very far in their studies. Learning material well is only determined on how hard you work at it, to understand it."

> "Yes, someone doesn't have high natural ability in physics or astronomy they can still put in the extra effort and succeed in the class. They just have to be committed to it."

Although not always the case elsewhere, within the written responses for this question hard work was also present in most statements which included mentions of resolve. Of course, the coding of "resolve" may be redundant given that students definitions of "hard work" often contained ideas of persistence and determination. The last prominent response involved discussions of student approaches to learning:

> "Yes, I think a person just has to look at a subject in a way that appeals to them the most. For example, if a person gets lost when being lectured on multiple formulas, it may help to look at the concepts behind them or see them in action through experiments."

> "Absolutely. With the right instruction and instructors, as well as the right frame of mind, I believe students can accomplish whatever they have their sights set on."

As seen in the examples shown above, these students base a portion of their response on factors which they believe can have an impact on their learning.

The clear and present theme for responses to this question was geared around hard work and its role in learning. At the heart of this theme is the likely belief that, regardless of the status of one's natural ability, hard work will enable you to learn the material. To elaborate further, hard work touches on principles of learning/studying and thus this question may be probing the efficacy of learning/studying strategies grouped with hard work and, to a lesser extent, in comparison to natural ability.

EBAPS Question 10

The question as given to students was:

> "It is often the case that a scientific principle or theory just doesn't make sense. In those cases you have to accept it and move on, because not everything in science is supposed to make sense. Do you agree or disagree with this viewpoint? Explain your thoughts."

This question is similar in composition to that of EBAPS question 10.

From a philosophical standpoint, this question (like its EBAPS counterpart) explores the structure of scientific knowledge. Formally, it explores the degree to which information within science is independent or a coherent whole. Analysis of EBAPS data indicates strong undertones of several other EBAPS axes. "Source of ability to learn" shows up with the presence of mindset (fixed vs. malleable), and "evolving knowledge" may be present with students displaying views of science as yielding either concrete knowledge or being lost in ambiguity. There also appears, to a lesser extent, to be aspects of "nature of learning and knowing". This appears in responses where students are seen to acknowledge the absence of personal knowledge (as it pertains to the question), how to create understanding, and/or what not understanding indicates. Lastly, several (if not all) of the axes discussed above could instead be a reflection of student views about science. Although these concerns would seem to indicate a fractured question, it does appear as though there exists a prominent theme within the responses and will be discussed further below. What is also seen within these responses is sophisticated insight within a response that agrees with the statement (the unsophisticated answer). For example:

> "I agree. For example, evolution didn't and still doesn't make sense to many people but we know for a fact that it is real and we evolved through millions of years. I believe many things become more clear for people as time goes by and we just need to be patient to understand why."

Of these kinds of responses, students tend to focus on some aspect of time and the need for patience. Perhaps these students are focusing on the idea that the theory has been put forth but not all the evidence has come in so as to create more certainty. Either way, some of these types of responses are focusing more on what it takes to give credibility to the theory as opposed to whether the theory is capable of being understood. This could come from the potential dichotomy of the question itself. That is, there are two key parts of the statement that students may be focusing on, either "[...] you have to accept it and move on [...]" or "[...] not everything in science is supposed to make sense." A focus on the former may invoke in students a view about how theories/principles function. In the case of the student above, they appear to have confused whether or not a theory should make sense with the current credibility of the theory itself (as dictated by evidence towards it). This duality to the statement could also explain the numerous aspects of epistemological beliefs that this question appears to elucidate.

In all, 47 students gave written responses to this question. Of those 47, 13 agreed with the statement, 4 were neutral, and 30 disagreed. More codes than is typical were assigned to aspects of these students' written responses. As a result, it may be best to discuss these codes in groups by similarity. The most prominent category is one which has previously shown up, student views about science. The least informative code within this category was that of the difficulty of science:

> "I kind of agree with this because some science is so intense that amateurs might not understand it compared to scientist or higher degreed."

Here, students allude to science as being an intrinsically difficult subject. There is little evidence regarding why they hold this belief and if this view reflects a deeper held belief about mindset (fixed vs. malleable). The next code pertained to the idea that knowledge within science accumulates gradually:

> "No, I disagree. There is always something new being discovered in science."

> "[. . .] More research is often all it takes to explain such inexplicable theories and principles."

The general response here focuses on the idea that information in science is being continually updated. It is possible that this belief could be simultaneously outlining sophisticated (or unsophisticated) views regarding the tentativeness of scientific knowledge. That is, an understanding of how science works may lead to students appreciating the importance of evidence and reasoning regarding this evidence, but could also cause students to view scientific knowledge as unstable. Keeping with the trend of tentativeness of scientific knowledge comes the next student idea, that knowledge in science is meant to be challenged:

> "Sort of agree, again, like with 1, they probably know more than me but if there is valid evidence against that theory then that is the basis of proving it wrong."

> "I disagree, science is something that can be argued and should be constantly tested and expanded on. Basically not everything is known so it is important to study the unknown to better understand the world."

Again, these students understand that knowledge within science is not set in stone. Furthermore, it is a staple of science that this knowledge be occasionally challenged for continued validity. The next aspect seen in some student statements is vague but has strong presence nevertheless, that science (and theories) are supposed to make sense.

> "I disagree with this viewpoint because most of the time the theory just don't make sense to the one student or else it wouldn't be a theory and that student just needs to work harder to grasp the concept."

> "Actually, it is the opposite. Everything in science has an explanation because it is, by definition, scientific [. . .]."

> "While I myself tend to do the same thing as the sample students in question 3, I do think that one needs to understand the theories and principles. Science is supposed to be a system of logic and make sense."

It is uncertain what these views are an indication of. To expand upon this, perhaps these views are simply the student showing a sort of faith in science, without any deeper understanding of how science works. On the other hand, students may be displaying this belief because they truly understand the nature of science in general and can appreciate the potential value of information that is presented in science. To make an analogy, anyone who views the Mona Lisa may simply appreciate the greatness due to the renowned status it has achieved. However, an art connoisseur viewing the Mona Lisa may hold a unique understanding for what makes the piece of art unique, from composition (inclusion of background in portrait) to painting technique (Sfumato). The last code in this category was one which did not fit a general idea and was thus simply associated with a variety of views about science and/or scientists:

> "I both agree and disagree with the viewpoint. I believe that some things in science really don't make sense but, we know them to be true because of the testing that has been done to prove it. Yet we don't completely understand why it is the way that it is. However, I disagree that you just have to accept it. I feel you only except it when all the fact are there to prove it, with little to no discrepancies."

> "Science doesn't always make sense. You may get an outcome you didn't anticipate. You may have variables that you didn't even know about until the end result. Sometimes there are components to theories or principals that are inexplicable because the components haven't been discovered or correlated to a specific thing before."

> "I disagree with this. I think that if scientists are going to make bold statements about things that are very important, there needs to be no doubt in their mind that they are correct. I want to know exactly what the theory is about and what is means before I just believe it."

Key aspects of these statements simply reveal general ideas regarding science and/or the role of scientists, with little further discussion. A relatively adequate pattern in these three examples above could be associated with the process of science. By this it is meant that these students are displaying views regarding how science acquires knowledge and/or how results are put forth to the public. It might be enlightening to further probe these students by asking a question such as "how do scientists gauge whether they have enough information/evidence to be comfortable with making a conclusion?" This type of question would help grasp what students reflect on when thinking about the process of science and the validity of the information presented. The next category discussed shares a lot with the EBAPS axis "nature of learning and knowing." The primary code within this category was that of obtaining an understanding:

> "I think it should make sense. I might just have to learn some more information before I can understand it."

> "I believe that if you don't understand a concept, then move on and ask a professor when you get the chance. It is easier to move on and get your work done and learn with help."

These responses generally display a philosophy about what should be done regarding the current state of their knowledge. Responses were too vague to indicate if how they intended to address their missing knowledge was through absorption of information or through more active mean such as working through material. Related to the previous code was a code which noted students who acknowledged that they could be missing out on knowledge if they just moved on:

> "If something does not make sense to you and you just decide to give up and accept it, then you have given up an opportunity to enhance your knowledge and learning."

This is potentially the first step in a student actively wanting to put effort towards learning the material. Not to imply that this step must occur, simply that these students are recognizing that there is value in the information not currently accessible to them. The last trend seen in parts of some student responses seemed to exhibit metacognitive ability:

> "I disagree. I'm not strong at memorization, so if I don't have a good understanding of some topic then I will not be able to apply it, and it will slip from my mind quickly. Lack of understanding in science is usually an indication that my fundamentals are weak, and I need to do some work on math or basic skills which are related to the topic."

These students contain a portion of their response which indicates monitoring of their understanding. Namely, they are reflecting on what it means to them personally if they do not understand the material. This goes beyond just the acknowledgment that they are missing out on information, as there is a clear indication of how knowledge and understanding relates to the self. Often times, multiple codes in this category will show up in a single statement:

> "If something that I am learning contradicts what I believe, then there is an opportunity to learn. If I spend some time examining my perceptions and previous experiences, I may find a persistent misunderstanding which I can correct. Alternatively, I may find a problem with what I am being taught."

The above student example shows this. "If something that I am learning contradicts what I believe, then there is an opportunity to learn" is the student acknowledging that they are potentially missing out on knowledge. "If I spend some time examining my perceptions and previous experiences, I may find a persistent misunderstanding which I can correct", is the student discussing a means by which they may go about constructing the knowledge that they have deemed to be absent. Combining "if something that I am learning contradicts what I believe [...]" with "[...] I may find a persistent misunderstanding which I can correct" seems to display metacognitive ability in both their monitoring of personal knowledge and their reflection upon what is might mean to not know. The final category is also linked to an axis of the EBAPS, the "source of ability to learn":

> "If something does not make sense to you, keep following it and asking about it until it does make sense. Everything that has ever been theorized has made sense to someone, why can't it have the possibility of also making sense to you?"

> "I neither agree nor disagree. There are times where accepting the prevailing theories is beneficial and allows people to not get bogged down. However, it's not because things are or aren't 'supposed to make sense;' all theories and principles should be capable of being understood by anyone with the will to do so."

As exemplified by the responses above, students here exhibit a kind of persistence in attempting to learn the theory/principle. What is unclear is whether this philosophy is, fundamentally, a manifestation of their mindset (efficacy of hard work, more specifically) or their views regarding science. It may be argued that the former example leans more towards a belief in malleable mindset ("[...] why can't it have the possibility of also making sense to you?") whereas the latter may be linked to beliefs about science in their giving science the benefit of the doubt ("There are times where accepting the prevailing theories is beneficial and allows people to not get bogged down. [...]"). The latter example is also fascinating in that it appears to piece together views about science with views regarding the efficacy of hard work ("there are times where accepting the prevailing theories is beneficial and allows people to not get bogged down. [...] all theories and principles should be capable of being understood by anyone with the will to do so"). The other code in this category is quite similar, the view that it (science/theory) makes sense to someone:

> “I do not agree completely agree with the viewpoint because someone, somewhere, somehow had to come up with and understand the scientific principle or theory or it would not be counted as such, and there are plenty of people who follow in their footsteps and research the principle or theory.”

> “I think that a scientific principle or theory should make sense although they may not make sense to everyone.”

Again, students here appear to be piece together views of science and views about mindset, with the result being that the scientific principle/theory should make sense to someone. It is not clear in these codes, however, if students believe they personally could make sense of it. The latter statement is intriguing in that the student may both be expressing a view about science along with a fixed mindset perspective. More information is needed regarding this claim, particularly what the student means by “[. . .] may not make sense to everyone.”

The prominent theme with this question seems focus around the student belief that “science is supposed to make sense” and hence this question appears to invoke student views regarding science. It might be said that this view not be possible without a belief in malleable mindset, however (note that the vast majority of students in other questions have indicated a belief in malleable mindset). Yet of those who disagree, nearly all of them display some code associated with a view about science. These views, as mentioned previously, are not necessarily always unsophisticated in their views about science.

EBAPS Question 11

The question as presented to the students was:

> "When handing in a physics or chemistry test, you can generally have a sense of how well you did even before talking about it with other students. Do you agree or disagree with this viewpoint? Explain your thoughts."

This question was intended to mimic EBAPS question 11, and essentially does not deviate from that question.

Since the EBAPS question and the question above are effectively the same, we expect that this question probes the "nature of knowing and learning". Essentially, that it is assessing aspects of student metacognition and whether they are active or passive learners. However, a sizeable number of students seem instead to reflect on study strategies and the efficacy of these study strategies (as will be seen below). Another concern regarding this question is the portion which states "[...] even before talking about it with other students." This segment is there to help isolate the student from considering the effects of an outside influence and thus get them personally monitoring their own understanding. Since so many students are focusing on study preparation prior to the exam in answering the question, the portion "[...] even before talking about it with other students" seems to function as a distractor instead. As such, a visible number of students are disagreeing with the statement because they believe discussing with other students to be beneficial toward understanding how they did. Thus, disagreeing for the reason stated above may reflect metacognitive ability, but it may not. Discussing this perspective further, a good number of these students view science exams as having right or wrong answers (little to know no ambiguity), and in exams from most any physics classroom these students would often be correct in this belief. Hence, exams themselves may function as a discussion with other students would, in that they reveal the answer to the students (i.e. they know it or they don't). To summarize, the exam is telling the student if they have the knowledge. It is less about the student actively monitoring the state of their own knowledge and more about it being told to them by the exam. Purely conjecture, but this question might not best elucidate metacognitive abilities in students because it takes place during an exam. If the question were instead to take place during the act of studying, this may better help bring forth the nature of knowing and learning. For example, the question could be "when studying for a physics or chemistry test, you can generally have a sense of how well you understand the material before discussing it with anyone else." That is not to say, however, that metacognitive practices cannot be activated during or after discussing material with others.

Of the 99 students who volunteered written responses, 4 didn't respond definitively (hadn't taken physics or chemistry), 4 both agreed and disagreed, 16 disagreed (unsophisticated), and 75 agreed (sophisticated). Oddly enough, two codes were frequent amongst the 20 students who did not fully agree with the statement but not among the students who did agree. The first was discussed above, that being students believe talking with other students to be beneficial in knowing how they did on an exam:

> "Disagree. I never know how well I am doing in any scientific activity without comparing results with peers."

With 5 of these instances occurring in the 16 who disagreed, but none of them in the students who were neutral (agreed and disagreed). There were 6 of this instance for the other 75 students (who responded sophisticatedly). The second code was that of discussing that they perform oppositely of how they believe they did:

> "Yes and No, If i feel good on a test I normally do bad and if I feel that I didn't do good I normally do alright on the test."

> "I almost always feel the opposite of how well I perform on test when turning them in. If I feel really good, I probably did very poorly and vice versa."

All 4 of the neutral students contained this view and 6 of the 16 students who disagreed contained this view. None of the students who agreed had this particular view. What this code says about student views in general is unclear. It is possible that these students are actually showing a poor ability to monitor their own understanding, and hence perform poorly on exams for which they've prepared. It could also reflect test anxiety within the student, further interviews would need to be done to elaborate in more detail. Two codes essentially share the top spot for being the most frequently occurring. The first, which was seen in 30 of the 99 responses, had to do with content familiarity:

> "Yes, I agree with this viewpoint because while taking a test you are able to create a pretty good understanding on your knowledge or lack of it on the material."

> "yes. you typically know what you know and do not know"

> "Yes, you can base how well you did in general on how many questions you had a definitive answer for."

The general trait here is awareness. Students claim they are aware when they do and do not have an understanding of the content. For an entire test, this view manifests itself in how many questions they did or did not have a comfortable understanding of. Again, it is unclear whether this perceived content familiarity reflects metacognitive ability. With confidence we may say that students are associating this view with their study preparation, as 10 out of 30 responses that were coded in this manner explicitly discussed study preparation.

> "I agree with this statement, because after you are given the test and you have taken it, you generally should be able to tell if you studied the right material or not. There have been several times where I have sat down to take a test and have realized right then and there that I studied different material."

This is a direct example of how students are reflecting on their study strategies when responding to this question. What is not known in these responses is if students are also reflecting on how they monitored their understanding (metacognition) during studying or during the exam. The other most frequently occurring code was the mention of study preparation, something discussed in hearty detail already:

> "Yes, generally if you studied or didn't study you know how well you did on a test. On the first test I didn't think I needed to study, I regretted this. Second test I knew that I did well o the multiple choice but not short answer and it showed in my results."

> "I can always tell how i did on a test based off of how well i prepared for it."

Over one-fourth of student responses (29 out of 99) contained a part which focused on their study habits prior to taking the exam. It is for this reason that this question may be doing a better job at drawing out student views regarding the efficacy of good study strategies as compared to their metacognitive ability. Another often-occurring view is rather simplistic, either you know it (the answer) or you don't:

> "I agree with this viewpoint becuase you either know or don't know the information tested on."

> "I agree because I know when I get things right and I do not in science."

This was made distinct from content familiarity because there is an aura of absoluteness within these particular student responses. Unfortunately, the nature of this dichotomy remains unclear. We can say that this behavior is occasionally a manifestation of how students prepared for the exam, with 7 instances out of 22 possible discussing their study preparation alongside this belief. However, this code could also reflect student views regarding science as being unambiguous:

> "I agree to a certain degree, science is often unambiguous. You either know the question, or you don't. It is difficult to bullshit your way through a science based test."

With this arrives another often-occurring belief, the belief that science in the classroom has definitive answers.

> “I agree because those subjects generally have specific formulas and processes that you follow and it is very obvious when I don’t know what I’m doing.”

> “I agree because science is nothing like say english; where you can write an entire essay and not have a clue you did anything wrong in it. But you will know instantly if before you even answer a question if you don’t know what you are doing you can’t beat around the bush and get a real answer with science or math.”

As would be expected, 12 of 13 responses which contained this belief also contained the belief that you either know the right answer or you do not. There is little insight indicating that these beliefs correspond to beliefs regarding the tentativeness of knowledge within the field of science. However, these views do appear to focus more explicitly on how science knowledge is tested within the classroom. Continuing with the theme of views regarding scientific knowledge in the classroom comes another code which represents students who refer to science problems as being chaotic:

> “I agree. I think there’s little mistakes (like math errors) that could trip you up. It’s mostly if you know the concepts or if you don’t know the concepts.”

> “Disagree, because those topics are so complicated there is a chance for error in other students work.”

One may have already deduced from the student responses above that by “chaotic” we refer to the idea that a miniscule error can manifest if an incorrect answer. It is possible that this view is also present when it comes to student ideas regarding the nature of scientific work (see written responses to EBAPS question 22), but no responses clearly indicate this. The last aspect of student responses that occurred often enough to be of note was that of “feeling”:

> “I agree that we have a good feeling about how we did on an exam because our intuition knows us best. Our gut and heart know things before our heads have a chance to catch up. This is not always 100% true, but the majority of the time I think our body knows.”

As seen in the example, students associate how well they did on an examine with an emotional state of being. This is a slightly vague category but many students mention feeling along with how prepared they were and/or how familiar they were with the material:

> “I agree with this viewpoint because if you study hard enough and long enough you should feel good about the material you are taking, i think that if you have no clue what’s on the exam you won’t do well and therefore feel really bad after.”

The overarching theme present within these responses is that of exam preparation. Before turning in a science exam, students typically associate how well they think they did with how good of a job they did in their preparation for the exam. This could also be tied in with student views about the efficacy of hard work, in general. Alongside this theme is a theme which regards the nature of scientific knowledge in the classroom as being unambiguous, that there exists a “right” answer to questions. This occasionally results in students thinking there is a particular way the question must be approached and/or that.

EBAPS Question 12

The question as posed to students was:

> "When learning science, people can understand the material better if they relate it to their own ideas. Do you agree or disagree with this viewpoint? Explain your thoughts."

This question is essentially the same as EBAPS question 12, with the obvious exception of the attempt to draw out student thought processes when answering the question.

There is one concern when it comes to students answering this question, and that is the nature of what it means to "understand". This could be said for any questions which contain the idea of "understanding". To understand for some may refer to the ability to remember what is taught, while for others it could be the ability to connect ideas between various concepts within taught content. Student responses to this question unfortunately do not noticeably help outline what the possible repercussions of these beliefs could be. On a similar note, because of the question similarity to EBAPS question 12, this question should explore the sophistication of student beliefs as they pertain to being either receivers or constructors of knowledge, along with metacognitive ability. In other words, we want to gauge how they justify their "understanding" (authority vs. self) and how they reflect and monitor that understanding (metacognition). As the question is phrased, student responses seem to indicate that they are answering in a manner which elucidates the efficacy of that study strategy (relating material to their own ideas). Having stated that, this question does not seem to draw upon the student consideration of where knowledge comes from. Only one written response (out of 40) explicitly focused on this possible aspect of the question:

> "It is usually beneficial for students to make their own connections, rather than just be told one, we should use the given examples and build our own understanding."

This is the only example in which a student responding to the question openly discussed the value of being told something versus creating their own knowledge. Since so few students do this, the question may be suffering due to the absence of an authority figure. For instance, if the question were phrased something like "your science professor compares a classroom concept to a personal experience of his/hers. However, when learning science, students can understand the material better if they relate it to their own ideas." This is merely an example (one close to that of EBAPS question 26), but this question could help better weigh the value students put on making personal connections as opposed to having the connections dictated to them. This question appears to do a better job of probing metacognition than it does the nature of knowledge (authority vs. self), and will be discussed further below.

This question saw relatively fewer written responses, with there being only 40 total. Of the 40 written responses, 1 student was neutral (agree and disagree), 1 student disagreed, and 38 students agreed. The most prominent code wasn't coded for its clarity, but rather because it was so vague. Students would often simply state something to the effect that the study/learning strategy from the question improves understanding and/or becomes easier to understand:

> "I agree that people can understand material better if they can relate it to their own ideas, because then it is not such a foreign concept."

> "I think that it is easier to understand if you are able to relate the information to your own experiences."

As touched on earlier, the nature of what these students mean by "understand" is elusive, shedding no light into their views of knowledge. Understanding could mean remembering what authority stated or it could mean taking in knowledge from an authority and integrating it in with prior knowledge, ideas, and/or experience. Another code like that of "understanding" pertains to students emphasizing the effect of relating material to their own ideas. These students go beyond the vague belief of (it makes understanding easier):

> "Agree, anything is remembered more clearly when one can relate to themselves in simple language."

> "I agree with this viewpoint. I think that applying personal experiences to learning is always helpful for students to remember material, as well as to really understand it and utilize this knowledge long-term."

As can be guessed from the student examples, the majority of these responses state the effect as being able to better recall the material (they focus on memory). Interestingly, the student from the latter example does delineate between remembering material and understanding it. Whether this reflects metacognitive sophistication (having acknowledged a deeper meaning to understanding) or otherwise would need further questioning of the student. Not all of responses discussing the effect of the learning strategy are about memorization however:

> "Yes because maybe if you see it from personal experience you can grasp the concept more quickly."

> "I agree. In my experience, this strategy seems to work well and emphasize the importance of subjects to people."

These examples show that this learning strategy could quicken the pace at which the material is learned or possibly assist in communicating the importance of science to the public. Another trait seen in these student responses, although not with nearly the frequency of the previous two discussed, was the idea that students should be careful when attempting to bridge the concept to their ideas:

> "Disagree because their ideas could be inaccurate and that could cause confusion."
>
> "I am partial to this viewpoint. There are benefits to both. In one way, it will help you remember better if you relate material to your own ideas, but you don't want to warp the information, or it will not be correct."

Students with this viewpoint display the belief that mistakes can be made when making analogies or incorporating new knowledge into previous knowledge or ideas. The one student who disagreed with the question (the former of the two above examples) did so for this reason.

Two primary themes show up within these student responses. The first has to do with the importance of prior knowledge and the influence it has on understanding new information:

> "[...] leads me to think that previous knowledge helps greatly in understanding future knowledge [...]"

However, even the theme above could possibly fit into the next theme, that being the efficacy of study strategies. Many students were of the opinion that this is a good study strategy and that it will improve learning within science. It may be inevitable that when discussing the efficacy of study strategies that you are also, to a large extent, exploring views regarding the efficacy of hard work as well. As a result, metacognitive ability (or an aspect of it, at least) is something which may be truly getting assessed here since students are reflecting on the success of a particular learning strategy.

<u>EBAPS Question 14</u>

The question as put forth to students was:

> "Understanding science is really important for people who design rockets, but not important for politicians. Do you agree or disagree with this viewpoint? Explain your thoughts."

The above question was designed to be similar to that of EBAPS question 14.

The biggest concern regarding student responses to this question pertains to the extent of scientific knowledge that a politician should have. A noticeable amount of student responses seemed to polarize the question and treat it as if a politician should have a grasp of rocket science (or science, in general) as well as the scientists themselves do. This treatment of the question was seen to bring forth unsophisticated responses (agreement) because, while this view would be an ideal occurrence, it is not realistic. As such, modifying the question in a manner that avoids bringing forth an extreme expectation could be more ideal. One option could be simply to remove the mention of scientists from the statement: "Understanding science is not important for politicians." However, presentation in this manner may shift focus and put more emphasis on the politician than it does on the belief that science applies outside its typical domain. To expand upon this concern, several student responses were already observed to be bringing forth their beliefs regarding politicians, specifically how knowing about science can influence their political career. As the question currently exists, it seems to do a commendable job of bringing forth student views about the real-life applicability of science, as intended for the EBAPS question.

There were 101 written student responses to this question, where 81 answered in a sophisticated manner (disagreed), 14 agreed, and 6 were unclear in their responses. The first prominent aspect to numerous student responses was the belief that politicians should know science because they influence policy:

> "Disagree! Science is just as equally important for people who design rockets and those who are politicians. Politicians have just as equal major impact on people's lives, so they should know about science just as much."

> "I disagree with this statement since politicians need to know what is going on in science in order to try and fix and see problems."

As can be seen, these students hold politicians as being in a position of power, capable of influencing the lives of the individuals they represent. These particular coded responses took two forms. The first form was that of the belief that policy influences science, hence politicians should have some amount of knowledge about science::

> "I do not agree. Science is in everything and we should all know and care about it. Politicians talk and have discussions about science like climate change and how much funding we should be giving science to explore more of the universe. I am an education major and I still think it is important to know about science."

> "No, because politicians regulate and occasionally fund scientific endeavors, they should have a significantly greater understanding of science."

The next form represented the inverse of the previous in that instead of politics influence science, science influences politics:

> "Science is more important for people who design rockets, but it is also important for politicians. For example, science plays a role in global warming and global warming plays a role in politics."

> "I disagree with this viewpoint. Understanding the basic science knowledge is important for people who design rockets as well as for politicians. Because some of the policies decisions are expected to be based on a large part of scientific conclusions."

Each of these two types of sub-codes were seen to occur with approximately the same frequency. Fundamentally, this code and its sub-codes appear to emphasize the influence that science can/should have. What this means is that students are reflecting more on the importance and/or applicability of science as compared to the importance of politicians.

The most frequently occurring belief that student exhibiteds was that of science as being vital to society. Like the previous code, this coded belief has two prominent aspects to it. The first sub-code was related students expressing the view of science as being informative:

> "I [dis]agree, science is important to all studies. In politics you need to understand certain topics such as climate change."

> "I disagree with this statement because I believe that it important for politicians to know about things like global warming and why scientists explore the universe. Everybody on Earth, especially those who can make a huge impact on the way people live, needs to understand science."

Students here simply express the belief that science is capable of yielding important information and is thus associated with valuable knowledge. Furthermore, many of these students seem to focus on a specific issue (such as climate change) when referring to the importance of science. Because of these views, it may be more accurate to describe this code as "science as (situationally) informative." This is emphasized because one cannot say for certain whether these individuals view the applicability/importance of science as far-reaching or topic specific. The other sub-code of science being vital to society was that which expressed science as a kind of must-have, fundamental knowledge:

> "Science is an important part of everyday life. Like how you get hot water in your morning shower is important to everyday life if it breaks."

> "I disagree with this statement, I believe that science is a piece of core knowledge that is important for the general population to understand regardless of their field of study."

> "I disagree because science is extremely important for everyone and all people on this planet need to understand it."

As can be seen, these students also emphasize the importance of science, but on a much broader, all-encompassing scale. Typical responses here contain ideas that science is everywhere and/or explains many of the things we observe around us, hence it is of great importance to politicians. Unlike some earlier response types there is usually little emphasis on the politician present here, instead focusing on the capabilities of science as an informative agent. Another view which revealed itself, but less frequently, was associated with student views about politicians:

> "I agree with this, politicians do not really need to know as much about science. A little knowledge would be good, but I think they already are kind of air heads so knowing science is not of the most importance to them."

> "Agree. Politicians mostly focus on what is occurring within our world and have no interest in space as it doesn't benefit their power."

> "I disagree with that. Politicians don't necessarily have to understand 'rocket science' but it would be beneficial to them to understand environmental science and other forms of social sciences. If politicians understand that this can help them with their campaigns and future bills they want to pass retaining to some social science field."

These views about politicians were coded as such because the students put focus on the politicians themselves, such as how science may be used for their personal benefit or how science may be treated due to the competence of politicians. Of the 11 total instances of this trait being observed, 6 of them were seen within the unsophisticated (agree) responses. The last code which will be discussed is that of the student view that politicians (or people, in general) do not need an expert-like understanding of science, merely a basic understanding:

> "No science is everywhere and everyone should have at least some knowledge about science whether you are a rocket builder or politician."

> "I agree that politicians do not need to have as an in depth understanding of scientific concepts as those who build rockets. But, politicians do need to have a base understanding, much like everyone else in society."

> "I agree to a certain extent. it is not as important for politicians to know and understand extreme scientific data but they should have a general understanding."

> "It is very important for politicians to understand science, maybe not to the extent that the scientists who design the rockets need to, but politicians need to understand the components that go into designing these technologies."

It is worth mentioning that 10 out of the 14 unsophisticated responses (agree) contained reasoning in which at least part of their written response was coded for one of the previous two codes discussed. A concern which was previously discussed, these students are treating the question as if the politician should have as strong a knowledge about science as the scientist themselves. As a result, students have focused on this within their argument.

The most prevalent theme within the written responses was that of the importance/influence of science. This overall theme aligns up well with the EBAPS axis that this question should represent, specifically the applicability of scientific knowledge outside the classroom/laboratory. From a nature of science perspective, one might say that the theme here pertains to the belief (or lack thereof) that science has personal and global implications. The impact of views expressing global implications was likely influenced by the timing involved in this responses. Responses which were requested at approximately the same time global warming was being discussed within class.

<u>EBAPS Question 15</u>

The question as posed to students was:

> "When solving problems, the key thing is knowing the methods for addressing each particular type of question. Understanding the 'big ideas' might be helpful for specially-written problems, but not for most regular problems. Do you agree or disagree with this viewpoint? Explain your thoughts."

This is intended to be similar to EBAPS question 15.

There were several concerns with how students responded to this question. The first concern pertains to the number of responses in which it was difficult to understand whether or not they agreed or disagreed with the statement:

> "I think knowing the root of the problem is important and helpful for all problem solving. If you don't know the roots of what you are solving for, it is easy to get off track."

> "When solving problems I think it is important to know the methods addressing a particular type of question and no matter the question it should be answered to the fullest extent possible if the answer is searched for."

There were 79 responses in all, 8 of which had to be categorized as unknown. Within these unknown responses, 3 explicitly displayed confusion about what was being asked:

> "I don't fully understand this question but I think I agree with it."

> "I do not understand the phrasing of this question, but immediately assigning a method to a 'big idea' sounds more like religion to me."

Furthermore, there are hints throughout all responses, not just unknown, that indicate a lack of clarity for this problem. For instance, 6 of the 24 unsophisticated responses (agreed) seem to have actually intended to disagree with the statement:

> "I do agree because when I do an outline for a paper with my own idea I always start out with my thesis about the big idea."

> "I agree, I believe it is best to fully understand the 'big idea' of things, and then apply that knowledge to regular situations/problems."

Although they may have simply been careless when expanding upon their thoughts, the possibility that they believed they should agree to the statement when instead they should have disagreed could not be ruled out. This possibility is further supported in knowing that behavior such as this has been a rarity throughout written responses to all questions. While there is no clear indication of what may be “wrong” with this question, there are signs throughout the student responses. The issues that will be brought up will be seen again within the coded responses below, but shall be discussed in some detail now. The first potential issue with the question is that it is pieced into two statements “when solving problems, the key thing is knowing the methods for addressing each particular type of question” and “understanding the ‘big ideas’ might be helpful for specially-written problems, but not for most regular problems.” Those who agreed tended to focus on the first statement, while those who disagreed tended to focus on the latter statement:

> “I both agree and disagree. I do think knowing ‘big ideas’ helps in all circumstances, but sometimes that isn’t enough.”

It is possible, as seen in the example above, that a student may agree with one aspect and disagree with the other, creating a lack of consistency. Although it creates discord, it can still be said that at least these students acknowledge the importance of grasping the “big idea”, and thus the question overall is working as intended by varying the sophistication of responses. Where the problem does appear to suffer is in the use of two key terms within it, those being “big ideas” and “methods”. The term “big ideas” may leave students with a bit too little of structure, leading to confusion about what is being asked:

> “I am not entirely sure what this question is asking because I don’t know what it means by ‘big ideas’. For problem solving, understanding everything about it can be beneficial.”

The issue with the word “methods” functions similarly to “big ideas” in that it creates a bit of ambiguity, but this ambiguity is complex. Within responses students seemed to take “methods” to mean either “algorithm” or as some form of alternative to “big ideas.” Students who view “methods” as “algorithm” have beliefs which are easy to classify as either sophisticated or unsophisticated:

> “Agree- The fact goes back to simple algebra and the method or order that you go through the problem with change the answer from time to time so if you stick the method you know you will come out with the right answer.”

The above student agrees, which is the unsophisticated response, and the content of his written review indicates such. However, students who answer with the idea that “methods” are some alternative to “big ideas” have responses that are more difficult to analyze:

> “I agree, because when looking at an overall subject, you might be able to answer most questions to a great length, but when asked specifics about a certain subtopic you might not be able to answer it fully or at all.”

These students are viewing the question as posing itself such that "methods" are now viewed as specific details about the problem (the counterpart to "big ideas"). This leads to many students agreeing with the statement because "big ideas" are, by nature, going to lack the specificity needed to solve the problem. One student goes into detail:

> "I agree and disagree. I enjoy programming and often need to translate real world problems into code. The 'big idea' is important to understand in order to reach the end goal and what is expected or desired from it. However, that actual writing of the program is broken into many different pieces that are connected to create the final product. As example, we may need write a calculator program. The 'big idea' may consist of functionality, possible additions at a later date, and the look and feel. Once this is known we can break the program up into smaller pieces. We will need a user interface of some kind, but this is a separate piece than the data and data manipulation. We need mathematical operations, which can be broken up into their own pieces. We will also need a way for the user interface and the data to communicate with each other."

This example response is quite sophisticated, yet has the student both agreeing and disagreeing to the statement. The student appears to treat "methods" as being the finite bits of information, all of which are guided by the "big idea". Response data indicates that this dichotomy regarding how students treat "methods" appears to be the most influential complication in student responses to this question.

In all, 79 written responses were submitted for this question. Of the 79, 8 were unknown or unclear in their response, 24 agreed, 41 disagreed, and 6 were neutral (agreed and disagreed). It is worth noting that 6 of the 24 who agreed likely meant to disagree with the problem, based on the content of their responses. The most frequently occurring response was vague, in that the students simply were of the belief that "big ideas" were helpful in essentially every situation:

> "disagree. Understanding the big ideas is always beneficial."

> "I disagree. I think that understanding the big picture is important when answering any type of question not just specially-written problems."

> "I disagree; it is always important to have the big picture in mind when working with 'regular' problems."

No general, and seldom any specific, trends were seen along with this view as the students' reasoning behind this stance was either superficial or lacking entirely. Rivaling the previous response frequency was the view that knowing the "big idea" was vital towards working the problem:

> "The big ideas to a problem is what helps you solve the problem easier especially written problems because it helps you understand what you are doing in that specific class/project and help you succeed in that situation I guess to speak."

Unlike the previous belief, this view had students elaborate a bit further on how and why the "big ideas" are valuable. One stance was that the "big idea" was something which could continuously guide the student through a problem:

> "Big ideas are important and can help a smart individual navigate the complexities of a lot of problems. Nuance should not be ignored and is key to deeper understanding."

Yet another perspective was that the "big idea" was useful in finding a way of starting the problem:

> "I disagree. I think understanding the 'big ideas' helps us understand each particular question. While we can't generalize questions, understanding the big ideas gives us a place to start when answering a question. That's why big ideas are formed, to help us by not having to come to the same conclusion over and over again."

> "I [dis]agree with this viewpoint because while problem solving you often have to figure out what works best for you and the situation and people look at the big idea in a certain situation, it is easier to figure out. Not all problems require technical methods to be solved."

> "In my opinion, understanding the 'big ideas' is a great way to begin a project. Understanding the broad concept and reasoning behind what you are doing and what the chances are that you will succeed in your study settles the mind to then focus on the details necessary to obtain a conclusion based on the methods for addressing the particulars of each type of question."

A final, notable view within this general belief was that big ideas provided a way of analyzing and/or revealing more intricate detail:

> "I think that understanding the big idea and then diving into a smaller part of that idea then applying it to the problem is key to solving most questions."

> "Yes, knowing the main ideas are helpful but for any type of problem. Working from the main problem and ideas can help analyze smaller ideas to further illustrate the main idea."

Another notable occurrence within student responses was the acknowledgement that knowing the "big ideas" was a crucial part of actually mastering the problem:

> "I disagree with this because if one TRULY understands the big picture, he or she can apply the concepts to almost any situation."
>
> "If you know the larger background you can usually come up with an answer to any question. No matter if it is specifically written or a regular problem."
>
> "I disagree because understanding the big ideas of questions is necessary to fully understand the question and answer."

In this sense, to "master" the problem means to have achieved an ability to work through most problems pertaining to the concept. Another student belief is one which was discussed in part earlier, that knowing the algorithm means getting the answer:

> "I believe the methods are important so you know how to approach a problem. This could lead to a solution. Understanding big ideas won't be able to lead to a methodically approached solution."
>
> "I disagree with this viewpoint. Understanding the 'big idea' in a problem to solve is necessary in order to fully solve it. Solely knowing the methods to answer a particular type of question will perhaps help you get a good grade but won't help you in the long-run."

The last code for these responses was associated with students focusing on minor details (instead of methodology) in that they believed "big ideas" to lack specific and valuable information:

> "When trying to solve a problem it is important to have a broad understanding of many underlying themes. Just understanding the big picture leads to oversimplified solutions to obviously complicated problems."
>
> "I agree because understanding the big idea can help with certain issues and problems but to truly understand something allows you to use it in everyday situations better."

The overall theme within this question is difficult to pinpoint, largely due to the complications outlined above. However, we may generalize these responses into an overall problem-solving sophistication theme. This sophistication is, it appears, in direct relation to student views about the nature of knowledge (the simplicity of knowledge, specifically). Although student interpretations of "methods" were blurred, it may still be said that they maintained relatively consistent views about the latter half of the question, “understanding the ‘big ideas’ might be helpful for specially-written problems, but not for most regular problems.” What this consistency allows us to conclude is that students who valued “big ideas” tended to allude to a view of scientific knowledge as being interconnected. Those who answered unsophisticatedly were seen to either view knowledge as interconnected or as isolated, depending in part on how they interpreted (or countered, one might say) the word “methods”.

EBAPS Question 16

The question as posed to the students is as follows:

> "Do you believe that everybody could learn to think more scientifically if they had enough time and they really wanted to? Explain your answer."

Essentially, this question remains identical to that of EBAPS question 16. The primary difference being the order in which parts of the question are presented.

As put forth by the creators of the EBAPS, a question such as this is meant to probe student ideas about the efficacy of good study strategies, with a strong association to mindset. After parsing the student written responses there was only one primary concern regarding the manner in which students address this question. This concern arises due to the ambiguity of the phrase "think more scientifically." To think more scientifically not only had different meaning to students within the responses, but would likely receive different responses from those formally trained in a field of science. Below are some examples of various interpretations of "thinking scientifically" from students:

> "Yes, science is step oriented and process based allowing anyone to partake."

> "Everyone should learn to think this way because I think there would be less mistakes in the world, because people would actually know why they were doing something in their job versus just doing something because they were told to do so."

As can be seen, the former response appears to view scientific thinking as a process, as an algorithm of sorts. The latter response tends to focus on the outcomes of thinking scientifically and the benefits provided. In both these cases, the question (specifically the phrase "think more scientifically") seems to bring out views regarding the nature of science itself, which deviates from the intended beliefs about learning which were sought by the question.

Of the 82 written responses from students, 77 of them agreed with the statement proposed (sophisticated). There were two key points that students focused on the most in their written responses. The first was on hard work ; The view that thinking scientifically is possible given that students put in the time and effort necessary:

> "Yes, anyone can learn whatever they want to with enough hard work."

> "Absolutley, anyone can do anything they want to, if they put enough time and effort into it."

As previously discussed, this code for "hard work" likely also involves an implicit view into students' beliefs about study/learning strategies. For instance, consider the following response:

> “Yes, you can get better with practice and studying at almost anything.”

This example is another type of response which was coded for hard work (“practice” and “studying” are too vague to be considered study/learning strategies), but almost certainly symbolizes particular study/learning strategies and their perceived efficacy with respect to being able to think more scientifically. The second most frequently coded response involved student focus on the concept of “want.” Despite the question being posed as being able to think scientifically given both enough time and desire, students were by far more concerned with the need for desire:

> “Yes, I think we are all capable of thinking scientifically, it just depends on how badly you want to think scientifically.”

> “Yes. The key idea here is that they want to. When people put their minds to it, they can do just about anything.”

The focus on the idea of “want” seems to go beyond mere agreement to the question itself, as seen in the examples above. These students view the want/desire to do something as a potential roadblock in students who may attempt to become more scientific thinkers. This is not to be confused with passion/interest in the subject matter, as other (less frequent) responses did indeed focus on that. Quite simply put, students here believe that anyone can learn to think more scientifically, but if they do not want to then it is not likely that they will do so. It is unclear in responses how “want” may be related to things such as interest or mindset. For example, one can have the mindset that doing dishes by the end of every day is beneficial because it keeps the house smelling fresh, but that does not imply that one wants to do dishes. Similarly, you can want to do dishes yet not be passionate or interested about the act of doing so. The motivators behind “wanting” to do something remain unclear. Two other factors seemed to play a part in student responses, but to a notably lesser extent than those mentioned above. The first is what we may refer to as ability in science. Portions of several student responses indicate a dependency on their views regarding proficiency within science and would often associate this view with different individuals exhibiting a shortcoming in their ability to learn science:

> “I believe that if someone really wanted to change the way they perceive science and how they apply it, they could absolutely teach themselves to. It may require more work and effort with some and less with others, but everyone is certainly capable.”

Here we see the idea laid bare, that becoming proficient in thinking scientifically has varying degrees of difficulty for various people. In the example above, it is not clear whether this view is based on a belief in natural ability or if it is something more simplistic such as an individuals background knowledge in science. However, some responses exhibiting this trait were more specific in where the belief stemmed from:

> "That is a hard question. It seems that some people do have an easier time with science related topics, but then again someone who isn't predisposed to understand science could get better with enough time and work put into it."

> "Maybe, I think STEM major brains and non-STEM brains work a little differently. If everyone could be equally good at science and english why don't more people double major in extremely different courses?"

Here, with the phrasing "predisposed to understand science" in the former example we are lead to believe that views regarding natural ability are present and influential. Similarly, the other example above effectively states a belief that some students are naturally more capable than others. The last factor was as prevalent as the previous and has been touched on in a discussion above, that being the influence of interest or passion toward the subject matter:

> "If they really wanted to yes, but some people aren't interested. Some people just don't love science because we all have different interests."

Again, much like "desire", a student's ability to think scientifically can be dictated by their personal interest in doing so. Students who expressed this belief did not seem to elaborate on what might generate and/or sustain interest. By association to other codes, it could be that interest is developed and/or sustained through familiarity with (or exposure to) the material:

> "Absolutely! Thinking scientifically is nothing more than taking curiosity and what one already knows to find out more. I think everyone thinks scientifically in respect to the things that are important to them."

The prevailing themes within these written responses focus on student attitudes toward learning (interest, desire, etc.) and their willingness to do the necessary work (whatever that may be) as being key components in a students' ability to learn the material. This question seems balanced on the impact that these attitudes have combined with the efficacy of study/learning strategies. This question may have done the most satisfactory job of those within the written responses in assessing "source of ability to learn" as presented on the EBAPS, as the question attempts to remove time and interest from the subject.

EBAPS Question 17

The question as posed to students was:

> "To understand chemistry and physics, the formulas (equations) are really the main thing; the other material is mostly to help you decide which equations to use in which situations. Do you agree or disagree with this viewpoint? Explain your thoughts."

This question is intended to be similar to EBAPS question 17.

There are many concerns regarding the manner in which students responded to this question. To elaborate, one of the key portions of the question "[...] which equations to use in which situations" appears to be a source of confusion. Several students were observed to disagree (sophisticated response type) with the statement (either in part or in whole), but rely on defending their disagreement with the idea that one also needs to know when/where to use the equation:

> "You need to know both. Just knowing the equations won't get you very far if you don't know when to use them in the correct situation."

> "I disagree because it is not always about the formulas. Knowing the material behind the formulas is the best way to help you understand what formula to use for a certain problem or experiment. The material is as important as the formula and it is good to know both really well."

This means that students who have unsophisticated views may be answering in a sophisticated manner (disagree) for unsophisticated reasons (you need to know both the formulas and when to apply them). Knowing when to apply an equation doesn't represent sophistication if the students themselves view the whole process of solving a problem as being algorithmic. The response behavior may be due to students viewing this question as composed of two opposing statements, that the main thing needed to understand physics and chemistry is either knowing the equations or that it's knowing which equations are to be used in which situation. The responses above, then, may have students at least partially disagreeing with the former portion of the question (that formulas are the "main thing") because they are agreeing in part with the latter portion of the question. This possibility is given further credibility with the knowledge that other examples (e.g. see question 10) have had similar issues with the presence of multiple statements within an EBAPS question. Consider that EBAPS question 17 is probing "structure of scientific knowledge" by evaluating the extent to which students place an emphasis on discrete, specific knowledge about equations and when to employ them. This is contrasted by the value students place on knowing concepts, which may govern how these equations are connected and derived. If the above suggestion is indeed the source of the problem, then the question should attempt to capture its intent in one all-inclusive sentence. A suggestion would be to consider framing the question as "to understand chemistry and physics the most important thing is knowing the equations (formulas) as well as knowing when and where to apply the equations." Of course, another solution would be simply to separate these two views into questions of their own. Another concern with how students responded has to do with context. Many, if not most, students considered understanding physics and chemistry as it pertains to the classroom. Students place additional value on the need to know how to solve numeric problems because that is how most traditional classrooms gauge student understanding. To further clarify, it is likely that numerous students read this question not as "to understand chemistry and physics [...]" but instead as "to solve numeric problems in chemistry and physics [...]." This issue of how students interpret what it means to "understand" is something not necessarily unique to this EBAPS question. Every question on the EBAPS that contains the word "understand" (or a derivative thereof) in the context of science is likely to invoke different interpretations, in this case it may create a dichotomy in which students reflect either on their ability to grasp concepts or their ability to solve numeric problems. Ideally, written responses from students about what it means to "understand" something would be of great value. Nevertheless, the context of this question tends to cause many students to view "understanding" as being strictly associated with their ability to solve numeric problems, likely due to the emphasis the question places on equations/formulas. There are several instances where students acknowledge this context issue:

> "It depends on your career. If you'll 'never' work directly with physics or chemistry again in your life, then I think knowing just the formulas in school is sufficient. If you DO work in a field that directly relates to chemistry or physics, it is absolutely better to fully understand the 'why' and not just the 'what'."

> "Yes from a trying to learn the material and pass a class standpoint I think that's how its taught. If you were researching, it would be different."

These student responses explicitly view "understanding" in the classroom as directly related to their ability to solve numeric problems (presumably, as they require the use of a formula). However, these students are also sophisticated enough to acknowledge within their response that other definitions of "understanding" exist. Given all this information, there should be some concern as to whether this question is measuring personal epistemologies or measuring classroom expectations.

There were 99 responses in all, 8 of which had to be categorized as unclear/unknown. Within these unknown responses, 7 explicitly stated that they refused comment because they had taken a limited number of physics or chemistry classes:

> "I'm not sure. I have never had to take a physics or chemistry test so I don't know the material."

> "I do not know, I have not taken enough chemistry or physics to answer this question."

Of the other 91 responses, 40 disagreed with the question statement, 35 agreed, and 16 were neutral (both agreed and disagreed). The most common code among students was that of supplementing a response with a link to either classroom experience and/or working problems:

> "I agree. After taking physics in high school, I learned that it really is all equations with different scenarios. We would use the same equations over and over, but with different circumstances."

> "I agree. Chemistry and physics formulas are the base for all the questions. You must know what main equation to use for each problem based on the context of the question."

> "Agree. The equations are the main point of these courses."

"Working problems" and "classroom experience" were coded as one because they both represent expectations and are closely related. "Classroom experience" reflects the expectations that these students have in which their proficiency in a subject is gauged by their ability to solve problems requiring the use of equations. "Working problems" reflect the expectations that students have in which solving a problem means using equations (as compared to, say, a problem which focuses on providing a conceptual description). It is of interest that these students discuss equation-necessary problems (numeric problems) despite the question itself never mentioning the word "problem", only "understand". This is one reason why there is a concern (as discussed further above) that students are displaying an expectation (not epistemological belief) that their "understanding" in this context refers to their ability to solve numeric problems. The majority of these types of responses were seen within the students who agreed (unsophisticated) with the question. When this code was present within those who were neutral or disagreed, the presence of an expectation is usually countered with an emphasis on the need for knowledge beyond just that pertaining to equations:

> "Yes and no. I feel like a lot of physics has to do with the math and an equation you have to use to get your answer. However, the information given or the answer you're trying to find can be in the material and this helps you determine the solution. But there can be information in there that is not given and you just simply have to know."

The next most prevalent coding pertained to the belief of equations and/or math as being fundamental, the vast majority of which are seen within the students who agree. Despite what one might expect, this code does not necessarily reflect purely unsophisticated beliefs. At worst, all that can be said is that this code represents students whose view is ambiguous (due to a lack of explanation):

> " Agree, I feel as though formulas and the math of those two subjects are really at the core."

or yet another reflection of student expectation/s:

> "I agree with this viewpoint. Because chemistry and physics basically rely on the calculation as well as the formulas (equations)."

Yet other responses under this code hint at sophistication with respect to the nature of science:

> "Formulas are the backbone of science. Most scientific action does occur in a way that is predictable. If it doesn't, it is more likely an indication we don't have the full picture and it is our understanding or process that is flawed, not the act of science itself. The other material is not irrelevant to the question at hand, but their proper installment into these formulas is their best usage according to our knowledge."

The above response could imply that formulas are a manifestation of the observations made regarding nature, a sophisticated view of physics. The student here could be considered unsophisticated, but only in how they are not openly acknowledging the importance of conceptual/structured knowledge in the development of formulas. The curiosity, then, is if this student views formulas as the foundation of scientific knowledge (unsophisticated), or if they view equations as being a manifestation of accumulated, structured knowledge regarding a phenomenon (sophisticated). Their reference toward equations as being the "backbone" of science does not explicitly indicate an unsophisticated view. However, it could be considered unsophisticated if this student places emphasis on learning equations over valuing the development of a conceptual framework. Yet other responses in this code actually refer to the importance of knowing formulas with respect to their learning:

> "I agree, formulas are very important and can help you find answers you need. Taking physics and chemistry in high school was a challenge, but being able to memorize formulas and equations helped me to better understand material and do better in the classes."

Again, expectations about what it takes to do well in a class show up in this response. However, this student refers to their knowledge of equations as being a stepping stone toward learning other material. Another common trait in student responses was the argument that you need to be able to actually use/apply the equation:

> "Agree with knowing the formulas but you also have to know the laws behind it and know the proper elements to put in the formula. Just knowing the formulas isn't enough to fully understand chemistry and physics."

> "I do not agree with this viewpoint. I believe that calculations are important, yes, but if you do not know the baselines of how to apply the formulas, how could anything get solved?"

> "No, you need a good deep understanding of the content of which you are studying to utilize any of the information or formulas."

These responses refer to knowledge beyond just knowing which equations to use in which scenarios, that you also will need to know how to use them. An analogy here would be that an individual might recognize the need to employ a jackhammer (memorization/algorithmic) to break up concrete, but still not know how to actually use a jackhammer (creation of knowledge from concept/context). Similar to the previous coding is the view that one also needs to know why the equation works and/or what it means:

> "I disagree; I think the equations are incredibly important, but concepts are key. Science becomes dramatically more difficult if you can't put the equations in the context of the 'big picture'."

> "I believe it is more important to understand why things work the way they do than just how to calculate things. You may know all the equations for everything but if you don't ask why you will never discover anything new which is the most important aspect of science."

Both this code (importance of understanding why an equation works) and the previous (importance of understanding how to use/apply an equation) were almost entirely exclusive to the students who disagreed with the question statement (sophisticated). The last code worth mentioning is that which involves students linking an aspect of their response to real-life, more specifically life outside the classroom:

> "The equations aren't the main thing. It's unfortunate that we are quizzed so heavily on the specifics, a general idea is usually okay. Because when you're out in the field actually testing these, you will not have lab conditions and the formulas. The formulas are only as good as your prior knowledge of the subject. If you don't know anything about the subject at hand then the formula is no good."

> "I think in order to do any job or activity you must understand the material and why you are doing it so you could be more effective at it."

> "I agree with this because a lot of what you learn are the formulas and then situations where you use them and why you would use them. Sadly, that way of teaching does not really prepare you for the real-world sandbox where there are so many interactions between several different things.

These types of responses were more frequent within students who disagreed or were neutral. The student above who agreed with the question statement could still be said to be displaying a sophisticated belief, as they have acknowledged their response as being biased towards their classroom experience.

There are a couple overarching themes associated with student responses to this question. The first involves student views about what is expected of them within the classroom to perform well. This typically seems to manifest in the belief that they need to know how to utilize formulas in order to solve numeric problems in class. Another theme is students treating knowledge about formulas/equations as being indicators that you have some form of subject mastery. It is difficult to say whether this is exclusively an unsophisticated view, as in several cases students have taken this stance due to their views about what it takes to have created an equation within science (or that science is concerned with the application of math toward nature). The last prominent theme is one which could be called "problem sophistication". This can take several forms, such as placing value in the conceptual aspects of a formula (how to use it, why it is being used, etc.). "Problem sophistication" may also represent the acknowledgement that one must still need to know how to interpret any results that are obtained, a heavily conceptual requirement.

EBAPS Question 21

The question as posed to students was:

> "Which factor do you believe is more important when it comes to being successful at most things in life, hard work or natural born ability? How much more important is one over the other?"

This question is designed to motivate similar responses as would EBAPS question 21.

There are two primary concerns regarding the manner in which students respond to this question, both of which involve interpretations of "natural ability". The first concern involves natural ability as either pertaining to academia or physical capability:

> "Hard work and natural born ability are equal in the essence of successfulness. Of course, natural abilities and natural talent can lead a person far in life, but hard work is the essence of life. Personally, I have natural born abilities in athletics in particular, as well as talents involved in academia."

> "[...] I believe that hard work is way more important to the working class over being able to jump or run faster."

The above are examples of students at least considering alternative definitions of natural ability as potentially relating to athleticism. This is a deviation from what would be ideal, that being natural ability as pertaining to cognitive ability. The second concern is really just a broader definition of "natural ability":

> "I think that while we're taught that hard work always trumps natural born ability, I would argue that this isn't always true. If we look at natural born ability as other factors than just talents, intelligence and extend it to race, religion, gender, physical handicaps, and wealth, then suddenly in my opinion, natural born ability is far more useful than hard work. You will have to work infinitely harder if everything doesn't go in your favor from birth than someone who is privileged enough to receive everything that they need."

Again, this student is considering "natural ability" as referring to a wide range of other traits that could be brought into the classroom. Both above examples represent possible causes of error within the accuracy of student responses. It could be best if, within questions probing natural ability, that some type of definition of "natural ability" be established. Written response data of "natural ability" do indeed indicate that students hold a wide variety of interpretations to "natural ability".

There were 200 total written responses to the question. Of the 200 total responses, 9 were neutral (hard work and natural ability of equal importance), 8 believed natural ability to be more important, and 183 believed hard work to be more important. A common trend among student written responses was the belief that hard work creates/demonstrates attributes/skills:

> "Hard work is more important as it demonstrates that a student can learn any material, regardless of topic, instead of having to rely on having a natural gift in one subject."

> "I believe that hard work will always help to show one's true colors. Dedication and compassion can only truly be seen through how much work and effort one puts into tasks, therefore I think that hard work is more important when it comes to being successful at most things in life."

> "I believe that hard work is more important than natural ability when it comes to being successful at most things in life. I believe hard work shows character, discipline, and ambition which can help drive you towards success. Natural born ability doesn't require any effort, which means there is no progress or growth. Success is something earned not given."

> "Hard work is much more important than natural born ability. Hard work allows you to gain skills whether or not you were born with them. Hard work is much more likely to bring you success in life than your natural born abilities."

In a general sense, these students believe that hard work builds and demonstrates character. Specifically, someone who is capable of hard work has shown that they are versatile, resilient, and dedicated to the task. There are factors beyond that of purely hard work which students attribute to success as well, or at the very least associate as being necessary alongside the ability to work hard:

> "I believe hard work is the key to being successful in life. I think that people with natural born ability may have an easier time doing something, but if someone is passionate enough about something they can accomplish anything they want to in life. Hard work is the key to most things but most they need motivation to want to work hard and when they lose their drive it can be hard to get it back. You have to want to work hard to do well."

> "I believe one of the most important factors at being successful is drive, passion and motivation. You aren't going to be a successful part of society without pushing herself to your full potential and working to reach your goal every single day. With drive and passion you will find yourself being more successful. You need a drive to understand things and find ways to approach situations."

> "I think that hard work is the most important thing when it comes to being successful at most things in life. Most things are able to be learned or overcome. If someone is really determined to do great things they will put in the hard work to see their dreams through."

> "Hard work easily. You don't need to be gifted naturally to have a good work ethic, maintain a positive attitude, or be willing to learn. I believe that those three things can lead to success and ultimately make you happier."

Often these aspects of students' responses refer explicitly to motivation, interest, and/or dedication as playing an important role along with hard work. Occasionally these traits are associated with mindset itself, such as a willingness to learn and a focus on emotional state, demonstrated by the last student example. The most often seen trait within student responses was represented by two codes. The two traits (codes) were the idea that natural ability has a limit:

> "I believe that hard work is the most important trait to have in order to be successful in life. Natural talent only gets you so far and sooner or later hard work will beat out natural talent."

> "Hard work is far more important t than natural born ability. Natural born ability can only get you so far in life but without the right work ethic the natural ability would go to waste. On the other hand hard work can make people capable and very good at almost anything they put their mind to."

and the idea that natural ability is hindered without work:

> "I believe that hard work is much more important than talent. There is a saying that I have heard many times in my life that says hard work beats talent when talent doesn't work hard. That being said if someone with talent also works very hard they are destined for great things. Without hard work though even the most talented people will eventually fail. I believe hard work is the most important factor in becoming good at a sport, getting good grades, or even doing well in a job."

> "I think that hard work is more important to success in life. While the importance of natural ability is incontestable, it won't get you anywhere if you aren't willing to try to improve. I think that the combination of natural born ability and hard work is the key to success."

> "Hard work is more important. One may have a natural born ability but it is nothing without hard work to maintain and grow that ability into something stronger and better."

The former idea, that natural ability is limited, represents a view that natural ability can only take you so far while the latter belief is perhaps more accurately stated as a belief that even if you possess some natural ability you will eventually need to rely on hard work to succeed. In many cases regarding the belief that natural ability is limited, it is also seen that those with natural ability will eventually need to work hard to succeed. The latter of these two beliefs was the most coded response type within the written responses. Although both beliefs are similar and likely do complement each other (or that these beliefs need to be together as a single belief) it was not assumed to be so. As an example, a student who believes in a fixed mindset could also believe that natural ability has a limit. As so very few students show a fixed mindset, it is difficult to make any comment regarding this. Of the students who believe hard work as viable (malleable mindset), there are no clear examples of them also believing natural ability as being a limitless construct. Furthermore, it is unclear whether students view hard work as something which extends beyond natural ability or if it is something that will complement an individual's natural ability. The last aspect of the written student responses to be discussed is the belief that an individual can catch up to and even surpass those with natural ability:

> "I think it is a little bit of both, I think that people most definitely have a natural born ability for certain things such as creativity, logic, etc. But I also think that if someone works harder to make up for a lack of ability they can easily match and/or surpass their naturally born competitors. I think work ethic is ultimately the most important factor, everyone is a genius at something, but they can't do anything with it if they don't work hard."

> "Hard work is far more important than natural born ability. It does not matter how gifted you are at birth, there is always room for improvement. Everyone can and should work hard to improve their knowledge and their talents. Someone who starts with nothing can reach the same level and even surpass someone with natural born ability with hard work. Success comes to those who work hard and try, not to those who rely on what they were born with."

> "I think that natural born ability definitely can give someone the edge over other people. However if you work hard at something you can achieve anything that someone smarter, or more talented can."

As can be seen, this belief also ties in nicely with the belief that natural ability has a limit and/or the idea that natural ability is limited without hard work. There exist several uncertainties within student responses and yet another one is how someone who works hard and has natural talent compares (successfulness) to someone with little natural talent but works hard. From analysis of the written responses, it seems that many students view natural ability as a step-up or head-start in life and not as something that continually benefits you. As such, an individual who works hard and has natural ability would likely be viewed as a least partially more successful than one who works hard but lacks much natural ability.

When attempting to link an overall theme to these responses one is tempted to group this under an axis from the EBAPS known as “source of ability to learn”. However, students within these responses do not have the efficacy of either study strategies or hard work at the focus of their argument, particularly as they pertain to learning. Rather, these responses are focused around mindset and almost solely that. Specifically, these students appear to be considering the influence of hard work as it directly compares to natural ability. One could argue that the efficacy of hard work is at least present, to which there would be little counterargument. However, these responses do not seem to place the efficacy of hard work within a learning context as has been done in previous instances. Instead, hard work here takes a broader role, encompassing success focused around any task in general and not just within learning/understanding. However one might view hard work, the theme here is how students view the importance of hard work as compared to the importance of natural ability.

<u>EBAPS Question 22</u>

The question as put forth to the students was:

> "Which factor do you believe is more important when it comes to being successful at science, hard work or natural born ability? How much more important is one over the other? Explain your answers."

The manner in which this question is presented is fundamentally different than that of EBAPS question 22, yet attempts to maintain the same interplay between hard work and natural ability. At the heart of both the above question and its EBAPS counterpart is the phrase "to be successful at science [...]". This question was postured in a different manner than that of the EBAPS so as to help draw out reasons pertaining to why students value one construct over the other. Hindsight would have that we also attempt to draw out precisely how it is that one construct was more important that the other.

Although majority of students responded to this question in a relatively satisfactory manner, there is one primary improvement regarding clarity that may be beneficial to make. This question is intentionally similar to EBAPS question 22 and as such should (could) be associated with the EBAPS axis "source of ability to learn". However, portions of several student responses indicate that this question may not be faithful to its probing of student views regarding the efficacy of good study strategies:

> "Science is more based upon hard work being science is an intense study that involves mass amounts of information collection. Hard work is much more important than natural born ability because hard work is what is required."

This may be due to students viewing hard work in the context of being successful in science, or possibly in general, as different. Using the above response as an example, this student appears to be viewing "hard work", as it pertains to success, being more laborious in nature. This is in opposition to other EBAPS questions along this axis which seem to be probing hard work in the context of learning science and hence as it relates to study/learning strategies. It is possible, then, that there is some tie-in with nature of science to the student responses in this question. If one simply changed "to be successful at science" to "to be successful at learning science" then the narrative seen in the response above might shift so as to focus more on study/learning strategies. The context in which students were utilizing "hard work" within the previous, yet similar, question (EBAPS question 21) is not clear and would need to be expanded upon in additional interview work. Alongside this concern above is also the use of the word "successful", as one student's version of success may be getting a passing grade in the course, whereas another student's version of success may be putting forth a new idea in science:

> "People with natural ability can skate by getting pretty good grades without a lot of studying. People who want to work hard and study a lot can get the grade, and get up to or surpass the level of people w/ natural ability."

> "I believe hard work is more important when it becomes to being successful at science. Natural born ability may help but without hard work there would be no successful scientists. Scientists have to work hard to test evidence to prove or disprove hypotheses."

These versions of success may be what drives the different views regarding hard work within this question. How one could restructure the question with respect to this concern is more difficult, as any definitions/substitutions of the word "successful" are going to need to be considered for EBAPS question 21 as well. As mentioned earlier, students also have different views regarding what "natural ability" is. It may improve clarity and consistency to state an explicit definition of "natural ability" in its place, even despite the use of "natural born ability" in this question and "inborn natural ability" on the EBAPS counterpart.

Of the 65 student responses, 57 of them were of the belief that hard work was more important. The most frequently coded response was associated with the student belief that hard work is required, regardless of what one has for natural ability:

> "Hard work is more important, even the most gifted scientists need to work very hard to get where they are today."

> "Hard work is much more important. Someone could be born with all the natural ability, but would never be any use unless they actually put work in."

> "I think hard work is more important than natural born ability. People can have all the talent in the world, but if they don't use it, or don't use it enough, then they are never going to get anywhere. Hard work is much more important, as people who work hard, whether they have a talent for something or not, can achieve great things."

These responses hint at the idea that students may view natural ability as something of a back-burner trait than a dominant front-runner. Codes discussed later on will further strengthen this proposal.

The next most frequently occurring tendency was the presence of student emphasis on either having or lacking natural ability. This type of code is a bit more generalized than other codes tend to be due to the lack of a common focus amongst the students. Of the many sub-codes which could be expanded on, two shall be discussed because they occurred the most often. The first was the association of natural ability with motivation and/or interest.

> "I definitely think there is something to be said for natural ability. I think natural ability sparks an interest, people like to do things that they are good at–for the most part."

Although this student favors the role of natural ability within their response, their reasoning for it isn't all that unsophisticated. In general, people may tend to show favoritism toward doing work that they are good at. To elaborate further on the above statement, responses to other questions have shown indications that interest in science can be tied in with prior familiarity. It is possible that this student is subconsciously associating natural ability with prior familiarity, which subsequently generated interest. Another sub-code focuses on the beliefs that having natural ability makes learning occur faster and/or having natural ability means less work:

> "I think it depends on how much natural ability you have. If you have a lot of natural ability it doesn't take as much work."

Precisely how natural ability allows for this isn't something that many students focused on. One of the only discussions of how natural ability is beneficial involved making connections between knowledge with ease:

> "[. . .] science will get further with hard work than someone who does not automatically connect the information."

However, even this statement is not as explicit as one would like, as it could be followed up further by asking what it means to "connect the information".

The next two codes which occurred often enough to be of note are potentially closely related, and thus will be presented side-by-side. These views are the belief that knowledge is malleable and the belief that natural ability has a limit:

> "Hard work as it can overcome any level of natural born ability. Hard work is much more important than natural born ability, because it can be changed, while natural ability cannot."

> "I feel as if hard work is more important when in the science field. You can always work hard to explore new findings and your natural born abilities only go so far."

As can be seen in the response above, these codes often appear in tandem with one another. Although they were coded separately, if these codes were combined the resulting code would be the most frequently occurring code present within student's responses to this question. However, there is no reason to assume that the belief that knowledge is malleable is always interwoven with the belief that natural ability has a limit. Further interview work would be needed to make such a statement.

For simplicity, the last two codes will be discussed alongside one another, as they have shown up in responses to previous questions (e.g. written response data to EBAPS question 21). These two being the impact of interest, and the resolve that a student has towards being successful:

> “I do not think that being successful at science depends most strongly on hard work or innate ability. I believe that scientific curiosity is the most significant factor influencing a person’s success in a scientific field. Innate ability is beneficial, hard work is necessary regardless of ability, but those with a strong sense of inquiry or passionate interest in a specific field are more likely to do better than those who lack such driving forces.”

> “Hard work. Anyone can conquer any subject with enough time and dedication.”

Whether these students are associating these constructs with learning within science is unclear, but is something that students have previously put value on. To clarify, the question at hand refers to what is needed to be “successful” at science, not what is needed to learn science. In the context of the question, additional interview work could be beneficial in attempting to understand how students view “success”.

In looking at the themes present with this question, there was seen a single predominant view, that hard work is more valuable than natural ability because natural ability has limitations and/or pitfalls. There are two fields of thought when it comes to natural ability and hard work. The first is that success via these two traits is much like a bucket of water. How much success you start with is akin to how much water you initially have in your bucket (dictated by natural ability). From here, water (success) can be added to the bucket (by doing hard work) or removed from the bucket (by not doing hard work). The other view in this theme is the belief that natural ability is something which will continuously assist you whenever you do hard work. This is distinct from the previous view in that the previous view seemed to treat natural ability as more of a head-start whereas this view treats it as a constant source of help. Whether this distinction is important or not has not revealed itself to the researcher. A factor which is unclear in all of this is student views regarding how much of an advantage an individual who has natural ability and does hard work has over those who only have hard work as their prominent trait. Knowing this may help determine the nature and legitimacy of these two perceived stances.

<u>EBAPS Question 24</u>

The question as posed to students was:

> "Who do you agree with in the following discussion and why do you agree with them?
>
> Alex: Science textbooks shouldn't treat each topic as a separate 'unit', because they're not really separate. I believe material in one chapter relates to material in other chapters and a good science textbook should show that.
>
> Taylor: But most of the time each chapter is about a different topic and those different topics dont always have much, if anything, to do with each other. The textbook should keep everything separate, instead of blending it all together."

This question is intended to probe student epistemologies in the same manner as does EBAPS question 24.

The primary concern with this question has to do with achieving consistency within student responses. These responses would seem to be centered on two of the five axes on the EBAPS. The first is "structure of scientific knowledge" (which EBAPS question 24 is claimed to belong):

> "Many scientific topics feed off of each other. In many cases, you have to understand one concept in order to understand other ones. In this scenario, I agree with Alex."

As the student response above indicates, this individual is exhibiting a sophisticated stance regarding the axis in mention. Specifically, their acknowledgement of how topics within science "feed" off of each other (i.e. how topics within science share a common influential structure). The other axis that appears to be present in influencing student responses is "nature of knowing and learning":

> " I agree with Taylor because it's easier for me to understand the material when its separate."

This student's response is demonstrating a sophistication along this axis in how the student is identifying potential learning benefits with respect to the manner in which material is presented (it remains possible, however, that they are not seeking connections to other material, which would be unsophisticated). Furthermore, this individual also appears to be personally reflecting on a particular learning style, and hence engaging in metacognitive processes. The presence of these two axes gives rise to the possibility that agreement with Taylor is the result of a potentially sophisticated belief with the "nature of knowing and learning" while agreement with Alex may be the result of a sophisticated belief along "structure of scientific knowledge". Hence, an inconsistency is present and it is likely that this question is unreliable as a measure of unidimensional epistemic sophistication. The second concern with this question is the role that expectations play. Some students have responded in a manner which may represent their personal experience with textbooks:

> "I do agree with them. Science textbooks separate the material being discussed so that's easier to understand. By doing this, in no way separates one chapter from another. Usually, at least from my experience, one chapter will lead on to the next with the same applicable learning."

The above student claiming " [. . .] at least from my experience, one chapter will lead on to the next with the same applicable learning" is an example of how a personal expectation could be playing a role in the answering of this question. One could make the claim that this previous experience is a reflection of their ability to identify connections within a text, but this is no guarantee as some texts explicitly make these connections for readers. This issue of past experience being relied upon for some students could at least be partially diminished if the portion of the statement by Taylor "but most of the time each chapter is about a different topic [. . .]" were removed. This removal is suggested because in this part of the statement Taylor is calling upon previous experience with textbooks in providing their counterargument. Instead of relying on past experience, it would be more ideal to present Taylor as purely exhibiting a personal belief as the focus of their counterargument.

This question saw relatively fewer written responses, with only 40 in total. Of the 40 written responses, 10 students were neutral (agree and disagree), 12 students agreed with Taylor (unsophisticated, in theory), and 18 students agreed with Alex (sophisticated). As a result, there were only four prominent traits which were observed within these written student responses. The first, most prominent trait was the belief that a separation of material will make the material easier to understand:

> "I am in the middle with this one. I think the chapters should be separate and cover each idea individually to not cause confusion. But I also think that there should be something or way of showing how everything in certain topics can relate to each other or go off of similar ideas."

> "I agree mostly with Alex. Often times topics relate to each other in books but they do need a clear definition so they do not get muddled together."

> "I agree with Taylor about splitting the textbooks up, though not quite for the same reason he says. Textbooks being split up into chapters is a good way to divide the content in a way that makes it easier to learn for the students."

As could be predicted, the majority of these response types come from students who agree with Taylor. Again (refer to above discussion), this agreement with Taylor demonstrates some sophistication along the EBAPS axis "nature of knowing and learning." The next most frequently occurring trend was the belief that chapters should be explicitly linked:

> "I believe both are correct. Although textbooks should relate topics a little better, some topics should not be mixed."

> "Alex is correct because holistic approaches are best when trying to integrate several themes."

These types of responses most often presented themselves as being philosophical in nature, with little elaboration on why that belief is held. There were several responses which demonstrated the possibility that linking concepts was beneficial to learning:

> "Both students have merit. A science textbook covers a variety of topics, so I agree separating them is a good idea, but ones that may cross each other could be combined or at least referenced to."

> "I agree with Alex. It is helpful to see how different topics relate to one another, and how they come together."

The use of the phrase "could be combined" may indicate the former student believes that at least some science topics do share common ground ("structure of scientific knowledge"). Similarly, the latter student's use of the word "helpful" might reveal some form of sophistication as it would pertain to the EBAPS axis "nature of knowing and learning". Unfortunately, there is not enough discussion by these students to be able to state with certainty that their response/s reflect some specific form of epistemic belief. There were, however, a handful of students within this category who did provide more intricate reasoning for their stance:

> "I agree mostly with Alex, although Taylor has a good point as well. I think that well-presented scientific information should make relevant connections between topics to support understanding and critical thinking rather than memorization that will later be forgotten."

This student's emphasis on the outcome of making connections " [...] well-presented scientific information should make relevant connections between topics to support understanding and critical thinking [...]" aligns up strongly with sophisticated views regarding "nature of knowing and learning". This is a good example of how certain epistemologies can depend on each other. In order for this student to value making these connections, it could argued that they first have to acknowledge the potential for connections to exist. The next prominent code was that of the view that links between concepts do exist:

> "I agree with Alex in this discussion. Material learned in one unit usually has something to do with the units that are ahead."

> "I agree with Alex because everything you learn can help explain other topics and help them make sense."

This code is quite similar to the previous, as for the most part the only difference is in the treatment of "should exist" and "do exist" when discussing connections. Nevertheless, this difference was accounted for because "should exist" could represent the possibility that a connection does not exist. At the least this distinction may only represent a slightly varying degree of epistemic sophistication. Half of the responses in this code were unclear in about whether the student believed connections exist because of personal textbook experience or if the student believed connections exist as a result of their epistemic beliefs regarding the structure of scientific knowledge. An example of this uncertainty is seen in the former of the two student responses directly above, where their reasoning cannot be said to be either expectation-based or epistemically-based. The latter of these two above responses represents the other half of the students in this coding, who do take an epistemic stance (for the student above, this stance would be best attributed to "nature of knowing and learning"). The last code to be discussed is one for which students explicitly state that knowledge within science is connected:

> "I agree with Alex because topics on science are integrated with one another."

> "I agree with Alex because science intertwines. Learning about stars (astronomy) also connects to Chemistry when we talk about what the star is made of. Science should be taught in a way that is connected."

As can be observed in the responses above, these students are undoubtedly displaying the belief that there exists a structured network of knowledge within the realm of science. It should not be of surprise by now to know that essentially all of the students exhibiting this view in their responses agreed with Alex.

The two clear and present themes within student responses are those which have been discussed time and again above, "structure of scientific knowledge" and "nature of knowing and learning". Students answering to these questions either show some form of belief that connections within science should/do exist or they show a belief that seeking connections within science will be beneficial to creating a better understanding of the material.

<u>EBAPS Question 25</u>

The question as posed to students was:

> "Consider the following conversation:
>
> John: My physics professor, Dr. Jensen, has just an incredible naturally brilliant scientific mind. It's always so enjoyable to talk science with her.
>
> Alex: It is a lot of fun to talk with her, but when it comes to being good at science, hard work is more important than 'natural ability'. Dr. Jensen is brilliant because she has worked really hard.
>
> John: Perhaps, but some people are just smarter at science than other people. Without natural ability, hard work won't get you anywhere in science!
>
> Discuss the extent to which you agree and/or disagree with both John and Alex. Be sure to explain your answers."

The question above is intended to be similar to EBAPS question 25, but also attempts to uncover reasoning behind student choices.

A concern regarding this question is similar to the concern regarding question six from spring 2016 (the EBAPS question 22 "recreation") that was put forth. Since the current question is so similar to EBAPS question 25, one could expect that this question also probes mindset and the efficacy of study strategies. Again, however, study strategies do not seem to be at the forefront of the students' minds when answering this question. Instead, students seem more focused on the interplay between hard work and natural ability, and the role that each play. Another concern is the final statement by John (Anna, on the EBAPS). This statement appears to contain two independent beliefs for which students may agree and disagree simultaneously. These statements are "[...] some people are just smarter at science than other people" and "without natural ability, hard work won't get you anywhere in science". Although a healthy number of students do seem to hold the former belief, the overwhelming majority do not hold a belief in the latter. Response options on the EBAPS version do allow for partial agreement with John (Anna) or Alex (Emily), however. Since mindset seems to be the focus of students in this question, it may be advantageous to define natural ability within context (as students do have different definitions of natural ability) and adjust both student statements. The idea being that these changes might better help determine the extent to which students view natural ability (as an innate trait) as having an impact on their learning. This could be done by removing "natural ability" within Alex's (Emily's) statement and presenting a position such as "[...] hard work is what's important, not something like being born more capable of learning science, if that's even a thing." Alongside this, John's (Anna's) statement could be made less absolute and state "[...] If you're not born more capable of learning science, then hard work won't get you far in science!" (an unideal double-negative is within this statement, but the spirit of the change is what the researcher is attempting to convey). These changes seek to provide a clearer distinction between the extremes (fixed vs. malleable mindset). Retaining the EBAPS optional response of there being a balance between the two statements would be ideal as well. These changes would have certainly allowed for more clarity and consistency within the student written responses and, likely, better delineated choices on the EBAPS (fixed mindset, growth mindset, and both). Of course, the latter bit is merely speculation based on interview data (much like the proposed statement changes, motivated by what will be discussed below). All these modifications to the questions are geared towards better extracting views regarding mindset. A completely new framing of the question would likely be needed to better elucidate beliefs regarding the efficacy of good study strategies alongside views about mindset.

There were 65 total written responses to the question above, 23 of which believe a balance between both students is needed, 41 of which believed Alex to be mostly correct, and only 1 of which believed John to be correct. There were two dominant aspects to student responses. The first of which was the belief that hard work is required, regardless of natural ability:

> "I believe both have valid arguments. John is right that it requires a certain level of natural ability but it also requires a massive amount of hard work."

> "I disagree with John, and agree with Alex. No professor earns a doctorate without hard work."

Of this code, only 2 instances out of the 23 possible were seen in agreeing with both Alex and John, whereas 14 instances of 41 possible were present in agreements with Alex. Responses containing this stance often fail to mention natural ability or, if they do happen to mention it, relegate it to place of little importance. The next dominant belief involved the malleability of achievement via hard work. Like the previous code above, this code addresses the ability to achieve by doing work:

> "They are both right in a way. Some people are just naturally gifted but that doesn't mean that if they didn't have this natural ability they couldn't use hard work to achieve this. That's just ridiculous because if you really try with hard work you really can be great at things. A natural ability doesn't mean that's the only way to become great at things such as science and what not."

> "I disagree with the idea that natural ability is needed for science. If it someone's passion to succeed, then if they work hard enough they can overcome somebody with a large natural ability."

Compared to the previous code, responses within this code focus more on the ability to succeed by employing hard work, and often express this with consideration of natural ability. 4 instances out of 23 possible were seen in responses agreeing with both Alex and John, whereas 13 instances out of 41 possible were observed in responses agreeing with primarily Alex. A response tendency which may simply be an aspect of the previous two codes is the belief that natural ability is limited, that one can only go so far with it:

> "I agree somewhat with both Alex and John. I do think that natural ability comes into play with being good at science but it is not going to let you slack off. Hard work is a big factor when mastering a subject. you can not just go with the flow without putting any work in because you will not get very far. So hard work and some natural ability work together to master the subject in science."

> "I completely disagree with John. Hard work is way more important in every field than natural ability. Any natural talent can only get you so far - it's hard work and practice that set people apart."

Since this view often shows up alongside the other two mentioned above, it is possible that this belief is typically held in tandem and is not necessarily an independent construct. That is, responses indicate that you wouldn't expect someone who believes natural ability to be more important than hard work to also believe that those with natural ability are limited in what they can achieve. Of course, statements such as these are intrinsically specialized toward the demographics of this study (college students). Responses to this question (as well as EBAPS questions 21 and 22) display a vast majority belief in malleable mindset. This indicates that those agreeing partially with John (Anna) in the question posed above are not doing so because of his statement "without natural ability, hard work won't get you anywhere in science!" Instead, students are likely agreeing with John because of the statement "[...] some people are just smarter at science than other people." One possibility then is that these questions are similar because they are both probing the varying degrees of influence that natural ability has in life, as compared to hard work. Responses indicate that at this stage in a student's life, any beliefs that you cannot improve in science through hard work are all but non-existent. Instead, there is now a struggle regarding how much of a role, if any, natural ability plays. Students who believe that natural ability is still a prominent factor complementing hard work tend to agree with both John (Anna) and Alex (Emily), whereas students who believe natural ability to take a back-seat role to hard work tend to agree with Alex (Emily). It is also worth noting that students may be viewing "smarter at science" to being "faster at doing science", given the definitions of "natural ability" previously discussed.

Hinted at in the work above, the dominant theme within this question is that of hard work vs. natural ability when it comes to learning (or being a "success" in science). Some individuals will largely dismiss natural ability while others will say it can play up to an equal part to that of hard work. No individuals state that you cannot improve your proficiency in science, however. Views about natural ability are present throughout. To make an analogy with art, some students view natural ability as everyone starting with the same tools, but those with natural ability have a canvas which is already partially painted. Others view natural ability as having different sized paint brushes. Those with natural ability can paint with tools capable of broad strokes, but tools capable of fine detailed strokes symbolize hard work. The idea here being that some students view natural ability as a head-start whereas others view it as a source of continuous, but possibly unnecessary, benefit. Either way, all the students are trying to paint the same picture, that is, all students are trying to learn something about science.

EBAPS Question 26

The question as posed to students was:

> "Discuss the extent to which you agree and/or disagree with both Justin and Dave. Be sure to explain your reasoning.
>
> Justin: When I'm learning science concepts for a test, I like to put things in my own words, so that they make sense to me.
>
> Dave: But putting things in your own words doesn't help you learn. The textbook was written by people who know science really well. You should learn things the way the textbook presents them."

This question is designed to motivate similar responses as would EBAPS question 26.

The primary concern within the student responses to this question are associated with an entanglement between learning and doing well on exams/tests. Specifically, students show hesitancy to place more value in personal interpretations because of the need to know key terms and ideas on tests. Due to this, some lack of sophistication regarding making personal connections to the material may not represent epistemological stances.

Of the 191 total written responses to the question, 32 were neutral (agree/disagree with both Justin and Dave), 9 agreed with Dave (unsophisticated), and 150 agreed with Justin. The most frequently occurring idea among students was the belief that the process of translating information is a viable method for learning:

> "I agree with Justin because when you put things in your own words it helps you understand the material better and actually helps you encode the material better into long term memory."

> "I agree with Justin and Disagree with Dave, putting things into your own word at first is a good way to understand what they are talking about, and then once you understand that it makes it easier to understand what they are trying to explain in scientific notation"

> "Putting things into your own words does not mean the textbook is wrong. It means that you put it into words that you can understand and memorize easier."

All these students believe there to be a distinct learning advantage when it comes to putting things into your own words. What is unclear is the nature of how these students view learning. Many students use either the idea that it helps you better memorize the material while other students refer to the idea that it helps to better understand the material. Of course, it is also unclear what students mean by "understand", but it generally would allude to a deeper level of learning than simply memorization. A pattern similar to that above involved the belief among students that putting things into their own words is an indication of subject proficiency and/or mastery:

> "I disagree with Dave because in-order to understand things you don't have to memorize word by word. I absolute agree with Justin. I think it's super important to be able to put things in your own word. because that shows how much you actually understand what you have read."

> "I agree with Justin because if you can put the correct idea in your own words than you have a mastery of that concept. Because it's not short-term memory, you'll remember it for a long time."

> "I agree with Justin, I believe taking things from the text book and putting it in your own words means your actually learning it."

> "I agree that learning things the way the textbook presents it to start with is good. But, I know that i truly understand a subject when I can explain a concept correctly in my own words."

These individuals are showing a level of sophistication in recognizing that the ability to put something into their own words is associated with a certain level of understanding. This demonstrates a degree of metacognitive ability within these students, given their ability to monitor their own understanding. Another common aspect of some student responses involved the importance preserving information:

> "Both are correct, if you put everything into your own words, knowledge and meaning can be lost. However, if you can't understand what is written, it does you no good anyways."

> "I agree with Justin. I disagree with Dave. As long as Justin can correctly translate things into his own words, without anything being lost, it is effective for him."

> "I agree with Justin to some extent. Putting things in my own words can help learn the material, instead of just trying to memorize things, but if I miss the point I could be learning the wrong thing. So I agree with both."

These students are generally okay with what Justin proposed, with the caveat that no valuable insight is lost when putting concepts and ideas into their own words. These responses seem to represent a degree of sophistication beyond just recognizing that putting things into their own words as being beneficial. However, their hesitancy and concern for maintaining an accurate translation could be tied into performing well as opposed to learning:

> "[...] some words of advice for Justin would be to make sure that the concepts that he is writing into his own words don't miss the actual point of the question or statement that is being asked/said because the concepts are taken from the book and will be pretty close to the ones on the test. Make sure you know what you are writing and that it is similar to the original question."

Many students within these responses refer to the idea of memorization as being an issue:

> "I agree with Justin because although the textbook is fully correct, it is easier to remember these things through your own words and how you understand it best rather than memorizing the textbook definitions word for word."

> "I agree with Justin because the best way to learn something is to interpret it into something that you can understand. Just memorizing book definitions for a test will not help you to actually fully understand the concept."

> "I agree with Justin. Memorizing somebody else's words like Dave won't help you grasp the concept, you'll just be memorizing words. By putting things in your own words, like Justin, you're transcribing the information into a way that makes sense to you and therefore it will become easier to remember the concept and how it works."

Although these students put an emphasis on learning, they also focused on the disadvantages of simply memorizing the course material as presented in the text. While this code could be grouped with the code which represented translating as a method for learning, the discussion of rote memorization distinguished it. Students here are putting forth a view which indicates that rote memorization of the text is not as easy as memorization and/or understanding with your own words. The last code presented was also one of the most frequently occurring, the belief that people learn in their own way:

> "I definitely believe with Justin on this argument. Although Dave is correct that the scientific definitions are perhaps more accurate, I believe it is important to create your own definitions to understand most scientific concepts. For me, Science is very difficult to grasp, the definitions are complicated! It is important to find your own way to understand them to better understand the general idea of scientific concepts."

> "I agree with Justin's method, text books teach in way that keeps students from exploring. They can take creativity and exploration out of life, and if somebody can learn in a better way by own explanation then so be it. It is their technique. Dave is a follower, he does what is socially acceptable and doesn't branch out. I think everybody could be better off by creating and understanding in their own ways."

> "Both are technically correct to their own extent. In order to understand a topic you can learn it in a way that makes sense as long as the facts stay true and info isn't made up."

> "Justin: I definitely don't do this, but if it works for some people that's okay too. Dave: I don't think Dave is right either. I think it depends on how you learn personally. I do copy exactly what the powerpoint says or what the book says because I understand it best that way and so I'll learn the proper terminology for the tests."

These students are okay with Justin and his approach because they hold the belief that everybody learns in their own way. They acknowledge that what may or may not work for them may work for someone else.

The overarching theme seen within these responses would appear to be a hybrid of "nature of knowing" and "source of ability to learn", as seen on the EBAPS. Metacognitive abilities and the efficacy of learning methods appear to culminate in views regarding the value that students place on making personal connections to the material. The influence of desiring to get a good grade is certainly present within at least some student responses, but the extent of that influence is unclear. For example, some students place the desire to perform well on tests over the desire to make personal connections to the material.

EBAPS Question 27

The question as posed to students was:

> "Discuss the extent to which you agree and/or disagree with both Julia and Carla. Be sure to explain your reasoning.
>
> Julia: I like the way science explains things I see in the real world.
>
> Carla: I know that's what we're 'supposed' to think, and it's true for many things. But let's face it, the science that explains things we do in lab at school can't really explain earthquakes, for instance. Scientific laws work well in some situations but not in most situations.
>
> Julia: I still think science applies to almost all real-world experiences. If we can't figure out how, it's because the stuff is very complicated, or because we don't know enough science yet."

This question is designed to motivate similar responses to that of EBAPS question 27.

The only discernable concern regarding the manner in which students responded to this question was the example which was utilized. There were several instances in which students would agree with Julia because they are aware that science can explain things such as earthquakes:

> "I obviously agree with Julia in this scenario - there have been countless scientific theories on just about everything brought forward, tested, and supported. Carla is clearly wrong on the other hand, as science can explain, in great depth, earthquakes."

It may be in the best interest of problem clarity to provide an example that is a bit more abstract than earthquakes, such as the placebo effect, how the pyramids were built, or why the same face of the moon is always oriented toward Earth.

Of the 173 total written responses to the question, 13 were neutral (agree and disagree), 6 agreed with Carla (unsophisticated), 151 agreed with Julia (sophisticated), and 3 were unclassified due to a lack of clarity within the responses. There was not seen a particular aspect of student responses which stood out from the rest. That being said, the first trend within student responses that is to be discussed is the belief that scientific progress is limited by technology or some other form of external mechanism:

> "I agree with Julia and I disagree with Carla. Mostly, because we know that it's a lot of stuff that we don't know yet and usually it is because we do not have the technology yet or we haven't figured how to do it in general."

> "I agree with Julia because I think that science applies to almost all real world experiences and if we don't have the answer to something, then we probably don't have the technology to figure it out."

> "I agree with Julia, science can explain most things and the things it cant are limited by our knowledge."

As can be seen, these students certainly view science as applicable to the world and where science fails the students typically point at technology or human limitations as being the cause of a lack of explanatory power. There existed a similar response pattern that did not focus on why science is limited, just that it will eventually provide answers:

> "I agree with Julia. I believe that everything can be explained and that there is an answer for everything. Just because we don't know the answer right away doesn't mean that there isn't one. All that means is that we haven't found the answer yet."

> "I agree with Julia because I think science does explain most of the things that we see. The few things that we see that can't be explained with the laws of science I believe science will explain in the future."

These students provide no clear reasoning for their stance and as such are viewed as students who simply have a faith in science in that they believe science will eventually provide an answer. There are responses which belong to both the previously mentioned aspects, and provide some insight:

> "This is a very controversial conversation, I can see why people would agree and disagree with this. Personally, I can see both sides because I can see the fascination of learning science and putting it in the real world, but with this I can see how when learning this its not realistic on how we learn it. Julia's last statement I can for sure agree with more than disagree, a lot of things we don't understand is because we don't have enough evidence on."

> "I agree with Julia. While science can't necessarily explain every mystery of the world right now, it is mostly because people (or more specifically, scientists) haven't figured it out yet, and might be missing a key piece of information."

The tone behind these responses also has a touch of faith-in-science as well, but does provide some form of reasoning. In may be best, then, to say that students responding in any of these previously discussed manners may be exhibiting sophistication in the real-life applicability of science but having varying degrees of sophistication regarding their reasoning ability. A reasoning ability which may be at least partially based on their understanding of the nature of science. The belief that science applies and/or explains almost all real-world experiences was itself made a code:

> "I disagree with Carla. She doesn't understand that everything can be related back to science. Everything that we see, touch, eat, etc. can be related back to science in one way or another. I agree with Julia because she stated that if it isn't already known it is because we don't know it yet or it is out of our scope of knowledge. The human race will continue to gain knowledge to adapt to their surroundings."

> "I agree with Julia, I do believe that science does explains a lot of things that go on in the universe and the world. Especially, when you can predict certain outcomes in nature."

> "I think that science is in all things and that it is a helpful way to explain all the things in the world."

Although many responses (beyond those put forth thus far) could be said to have some form of sophistication regarding the applicability of science, these students focus at least part of their responses on this belief alone. The following two codes to be discussed could be grouped under nature of science or structure of science, as they focus on beliefs about how science behaves/functions. First consider the view that science is a problem-solving tool for understanding life:

> "I agree with Julia because the scientific process was created to help us understand natural phenomena that we don't yet understand, and applies to everything we can observe."

> "I think Julia is mostly right. Science uses the scientific method to understand and explain most visible or invisible things and ideas."

> "I agree. I believe that science is a continuous process that will never stop. What we don't know is all a mission for scientists to use science to figure out how and why things are."

> "I would have to agree with Julia because just because you don't understand something doesn't mean it doesn't follow some sort of scientific rules. Everything can be explained by science it is whether or not we understand or conceptualize the science."

The emphasis with these students is on science as being a process by which to acquire information. That science is a protocol which will determine understanding, if you do science you will get answers. Further detail regarding the specifics of this scientific process is not elaborated on within responses. In the first example, the scientific process is explicitly referred to as the scientific method, yet any intricacies involving the method are not discussed (e.g. the replicability of data). The other belief pertaining to how science behaves/functions is the reliability of science:

> "I agree with Julia when she states that science helps us understand almost all real-world experiences but Carla is also correct when she states that we don't understand all things. Science helps us understand things using evidence, theories, and hypotheses but we do not know 'everything' yet."

> "I think Carla's way of thinking is very dangerous in modern day society. That just because there are a few holes and gaps in science then we should replace all the things we don't understand with something else. Which is dangerous because that 'something else' probably isn't fact or evidence based thinking."

> "I agree with Julia and disagree with Clara. Science is backed by facts and professionals, and can explain most things in this world. However Clara is right on some level because not everything can be explained by modern science (YET)."

Similar to the previous code, one could say that these students are still focusing on the process of science. The distinction made is that the students here are focusing on more structure than those previously. Specifically, that science is something supported by and/or employing reasoning/evidence/professionals/etc. To summarize, the first code had focused on the process of science whereas the second code focuses on the credibility of science. Another trait of some student responses was the utilization of examples within their justification, most often involving history:

> "I agree with Julia and disagree with Carla because Julia states the truth because in the 1800s they didn't even know there could be such a thing as cars, but not we don't know life without them. 1950s people didn't know what cellphones and computers that we can't live without today. As we get smarter and understand more of what were given today we can continually improve what we are given, and that's why Julia's statement is correct."

> "I do not agree with Carla at all, i think that science has a explanation for everything it is just the fact of going out of your way to understand it. It sounds like Carla just does not want to try to understand the concepts of science in other applications. If she really wanted to do some digging she could find answers to a lot of questions about how the world works. Truly there are things that we do not understand yet but that will come with time, we as humans once believed that the earth was the center of the solar system and universe and with time science eventually explained why that was not true."

The use of examples was only employed by one student within the unsophisticated responses (although, there were only 6 unsophisticated responses in all):

> "I agree with Carla more than Julia on this thought, I do not believe that science can explain everything in life, most things in life are unexplored, can not be reached or have not been solved. Some examples of unanswered questions that god forbid not even science has answered, i.e. the placebo effect, as long as you truly believe a pill will have a certain affect on you it will, scientists cannot explain that mindset or physiological occurrence. Who built stone hedge, science cannot explain nor can anyone prove or figure out who or what built those. I think it is ignorant to think that science can explain all. If science could explain all there would be nothing ton question or discover, there would be no disease."

It is uncertain what this last student is leaning on when portraying these examples. They appear to be tying the domain of the applicability of science to personal religious views. This domain specificity is then justified through the use of examples that they themselves believe controversial (both the placebo effect and who built Stonehenge have some plausible explanations put forth by scientists). It could be that the employment of such examples may be a reflection of their understanding of the structure/nature of science. There exists another response trait in which students state that they believe science may never explain some things:

> "I disagree with Julia because I think that the reason we can't scientifically prove some things is not because we don't know enough science yet, it's because those thing aren't things that can be scientifically proven."

> "I kind of agree with both Julia and Carla. I think they both have really good ideas of their own. Depending on what how you are looking at their argument they both make sense. I feel like Carla is kind of exploring what lays beyond science. like supernatural phenomenon. where as Julia is very objective. in a sense they both are right."

> "Saying that we don't know enough science is like saying we can't prove the existence of a higher power because that knowledge has not been made known. Both, however, are correct. Science does indeed explain things seen in our world, but it does not explain everything."

> "I agree with Julia because I definitely think science explains just about everything in our life. There is always ongoing scientific research that we may never get the answer too but the fact that we can study it scientifically is huge. Everything from a simple conversation and interaction to a scientific experiment is science in the real-world."

It may not come as a surprise that 5 out of the 6 unsophisticated responses contained this belief, and 4 out the 13 neutral responses did as well. When this response was utilized by those who responded unsophisticatedly to the question, there was often little justification behind their belief that science encounters things in nature which may never be understood. Compare this to students who answered sophisticated to the question, these students typically lean on their understanding of the nature of science in discussing how science may not produce a well-agreed-upon answer (see final example above). The last code to be discussed was simply an acknowledgement among students that labs/school can be limited:

> “I agree with Julia. she understands the way science works. We didn’t know what those shining things in the sky were at one point in science. We have learned that they are stars. Carla is stupid. Of course the things you do in a lab don’t exactly explain how earthquakes happen. There is too much information to explain in just one lab. Scientists spend their entire lives to finding these things out.”

> “I agree with Julia, although I can sympathize with Carla’s statement and feelings toward science. We have all wondered, at times what the purpose of the material we learn is. For example, everyone sat in geometry wondering what we were ever going to do with all this seemingly useless information. But for me, as I have grown up and learned more and more, I have come to see that there really is a use for every scientific or mathematical principle.”

> “I agree with mostly Julia, because while we don’t apply most of what we learn in science in the classroom, this is largely because we haven’t learned enough yet.”

> “I do agree with Carla to a certain extent. A lot of the things I have learned in school about science is not relevant to my life or not very practical. Although I think the things I have learned are important I just don’t find them useful in my own life.”

Much like the previous example, those who had sophisticated responses associate the limitations of science in school as being due in some way to the nature of science. Specifically, those individuals often rely on the vast amount of information that would need to be taught in order to completely understand the phenomena. Only one individual in the unsophisticated responses displayed this belief, and it has been shown as one of the examples above. In that one instance, the student appears to lack sophisticated views regarding the applicability of science to real life.

There are many aspects to the student responses for this question which make it difficult to pinpoint a general theme. In most cases it seems that student views/beliefs about science play some kind of a role in their responses. There is also a heavy undertone of the applicability of science to real life, although there were instances where this applicability seemed dictated by their views about science. One could also say that the reasoning ability of these students is being probed. Specifically, in how they are justifying the connection between what is seen in school and what is seen in life.

EBAPS Question 28

The question as posed to students was:

> "Who do you agree with in the following discussion and why do you agree with them? Do you believe both make good points? If so, explain.
>
> Erin: Some scientists think the dinosaurs died out because of volcanic eruptions, and others think they died out because an asteroid hit the Earth. Why can't the scientists agree?
>
> Anna: Maybe the evidence supports both theories. There's often more than one way to interpret the facts. So we have to figure out what the facts mean.
>
> Erin: I'm not so sure. In stuff like personal relationships or poetry, things can be ambiguous. But in science, the facts speak for themselves."

This question is intended to be similar to EBAPS question 28.

In analyzing student written responses, there were effectively no concerns. Unlike in other questions, there was not a heavy emphasis on the example utilized. This question did an acceptable job on maintaining student focus on the overall statement and not on other distractors such as opinions on extinction theories.

Of all 77 written student responses, 62 agreed with Anna (sophisticated), 5 agreed with Erin (unsophisticated) and 10 were neutral (favored aspects of both Anna and Erin). The two most prominent beliefs expressed by students are not only beliefs that you expect them to portray, but beliefs which are also quite similar in nature. Those being the belief that evidence can support multiple theories:

> "I agree with Anna because while, we are given facts, these facts could support many theories."

> "They're both correct to some extent. Erin is correct that a fact can't be ambiguous because it is just that, fact. Anna is correct because both events could have happened but with the evidence provided, they are equal in what could be fact."

and the belief that different interpretations are possible:

> "I find myself to agree more so with Anna. Her viewpoint on the facts of the scientific community and their interpretation aligns more with my beliefs. I believe that evidence and facts can be interpreted by different people and communities in different ways to mean different things. The interpretation of facts is relative and can change from person to person."

> "I can see both perspectives, the second person is right because there are many ways to interpret facts, and facts could lead to different conclusions, but also person one has a point that we feel like facts should be paramount and empirical."

These two perspectives were coded separately, despite their great similarities, largely due to the possibility that each could have deeper meaning and not because each display a different meaning. Those who tend to put a focus on what the evidence ("facts", as the EBAPS states) says could be indirectly revealing sophistication regarding the role that information plays in the creation of not only theories, but in the development of conclusions. Thus, alluding to their evidence-based reasoning ability. Students who instead place focus on interpreting the data may be placing more emphasis on the role of humans (scientists) within the interpretive process. This is supported by the fact that this particular belief was the only one of these two which was also seen to be coded alongside instances where students would discuss details about how scientists interpret data:

> "I agree with Anna, because there can be multiple ways to interpret information, and some scientist, when doing research are just looking for a certain thing, but the data could hold multiple important things."

The example above is one which displays that this belief may be emphasizing the role that people play in the creation of theories and the development of conclusions. This acknowledgement of a difference between these two beliefs above does not imply that the students themselves are emphasizing information over personal interpretation or visa-versa, just that this observation is worthy of note. The next coded oddity was the stance that several students took in which they seemed to emphasize the importance of considering multiple perspectives/opinions:

> "I agree with Anna, because it is very open minded and considering of all the facts."

> "I agree with Anna. That's why scientists come up with different theories so that we can see the different thoughts and ideas to fully understand what could have happened."

> "I think each person is entitled to their own opinion and they both make good points. But I mostly agree with Anna. Data does have to be interpreted, just because you know the facts does not mean you know the whole story."

These students seem to place value on open-mindedness and the benefit of competing beliefs. Again, this could be another aspect of reasoning sophistication on display. Another prevalent aspect of student responses was the emphasis on time between present day and the past, as it pertains to certainty:

> "I agree with Anna. Even if you have facts and evidence that points towards one possibility, you can never know for absolute certainty what happened, unless you were there when it happened."

> "I agree with Anna. Both make valid points and I can see where they are coming from. The reason I choose to side with Anna is because when talking about things that happened very long ago you have to be willing to interpret things."

> "I agree with Anna. The evidence could support both theories, and there is no way to know for sure what happened unless we went back in time. There is often more than one way to interpret the facts."

In a more general sense these students are of the belief that certainty cannot be achieved without being present, that time is essentially obscuring the ability to make definitive conclusions. It is unclear, then, the extent to which they might value the information that is available. The trend of data-obscuring conclusions was also noted when students would explicitly point toward there simply not being enough data to establish a single dominant theory:

> "I agree with Anna because the facts that are available don't tell the whole story often."

> "I agree more with Anna in this situation. Theories often reflect the amount of information that is known about a certain topic. It's not uncommon for newly discovered things, or very ancient things, to have different theories regarding them because there is so little information."

> "I think they both have good points but I agree more with Anna because there are more ways to interpret the facts when nothing has been traced to find one particular conclusion."

The overall idea present within these students is that when one is lacking sufficient evidence, it creates room for ambiguity. Their view of theories is similar to the game Wheel of Fortune, where people guess at a possible phrase while only having a finite number of letters and context with which to work. This stance is more sophisticated than the previous, since more focus is placed on the role of evidence, whereas the previous stance put emphasis on desiring firsthand evidence. Another prominent response trait present was student views about science. There were several seen here, with the most frequently occurring being the tentative nature of scientific knowledge:

> "Once again, 'facts are the enemy of the truth'. Nothing is factual, right or wrong, in science. There is only 'most probable' and 'most likely'. Our realm of understanding is not capable of determining sheer truth, and that is why science is constantly evolving and reevaluating."

and the nature of scientific theory:

> "There can be facts to support both theirs but relationships are hard to determine. It is important to weigh the amount of evidence for each idea and search for consensus given a preponderance of evidence."

To encompass these views about science into a single nature of science aspect would be to say that these views pertain mostly to the philosophical tenets of the nature of science. The last code was one which has been touched upon earlier, views about scientist interpretations of data:

> "I believe that both of them have good viewpoints, because of the fact that science can be both fluctuate and precise. I agree mostly with Anna though, since there often is evidence that contradict each other. Different scientist might find different evidence to support their ideas. Despite the fact that one idea might be more believed by the masses than the other idea, still don't give efficient proof of one being true or not."

> "I agree with Anna more so than Erin. It's hard to understand something that doesn't have a definitive answer. And scientists will almost never all agree. Whether its 97% of scientists arguing climate change is man-made or 4 out of 5 dentists recommending a certain kind of toothpaste, it's hard to reach a total consensus."

There are aspects of these student responses which harbor beliefs about how scientists/people view data and/or how that manifests itself.

The theme connecting many of these student responses seems to be the sophistication of a student's ability to consider evidence and to reason accordingly. There are often heavy undertones of students views about science, but these seem to meld with and influence how students are reasoning as opposed to existing as an independent construct being tested. Given the relatively low number of students who either agree with Erin or were neutral, it is difficult to associate them with a theme. Instead, it shall simply be stated that the most repeated view amongst these students was that of facts as being rigid constructs which can dictate interpretation.

EBAPS Question 29

The question as posed to students was:

> "Who do you agree with in the following discussion and why do you agree with them.
>
> Jose: Science is like fashion; something that's 'in' one year can be 'out' the next. Scientists regularly change their theories back and forth.
>
> Miguel: Once experiments have been done and a theory has been made to test those experiments, the matter is pretty much settled. There's little room for argument."

This question is intended to probe student epistemologies in the same manner as does EBAPS question 29.

As this question was intended to bring forth the same types of responses as would EBAPS question 29 it should thus be testing "evolving knowledge", more simply described as the tentativeness of scientific knowledge. The question setup has Jose discussing an extreme in which the findings of science are little more than mere opinion (capable of changing on a whim) while Miguel is arguing for an extreme in which all scientific knowledge is set in stone. The reality of scientific knowledge is that it lay somewhere in-between these two extremes. The manner in which students responded to this question does not necessarily indicate anything critically wrong with the statement. The largest concern about the statement is the use of the word "theory" (or derivatives thereof). It can be argued that Miguel has a predominantly sophisticated response here, as theories are generally agreed upon aspects of the physical world. When Miguel claims "[...] the matter is pretty much settled" the use of the word "pretty much" betrays the absoluteness of his opinion and creates room for interpretation of "pretty much". When it comes to scientific theories, a generally agreed upon explanation that has been rigorously tested (i.e. a theory) could very easily be interpreted by anyone as something which is "pretty much" settled. Although, the subsequent claim by Miguel that "there's little room for argument" does further help allude to the extreme of science as being set in stone. To clarify a bit further, "pretty much settled" and "little room for argument" seems to cause students to treat the statement more like "it is highly probable that the theory is accurate." It could be of benefit, if the word "theory" must remain, to rephrase Miguel's claim as: "Once experiments have been done and a theory has been made to test those experiments, the matter is settled. There's no room for argument." This position is, undoubtedly, that of knowledge acquired via science as being absolute. The removal of the word "theory", could be even more ideal. Consider Miguel's statement rephrased as: "In general, scientists will make an observation, generate an explanation, then rigorously test that explanation for validity or lack thereof. Once these scientists put forth their conclusion then the matter is settled, with essentially no room for argument." This statement does not include the word "theory" and thus avoids any pre-held definitions of such that the students may hold (this appears to be present in student responses, see further below). Furthermore, the proposed statement would also convey a clear position that scientific conclusions (knowledge) are set in stone. In the end, both Jose and Miguel make relatively clear overall arguments, judging by the student responses. Purely conjecture, but students may be able to convey a clearer position regarding the tentativeness of science if these two arguments appeared as individual questions in a likert-style agree/disagree manner.

This question saw relatively fewer written responses, with only 42 in total. Of the 42 written responses, 10 students were neutral (agree and disagree equally with both Miguel and Jose ; sophisticated), 18 students agreed with Jose (unsophisticated), and 14 students agreed with Miguel (unsophisticated). There were prominent codes within the responses, but all these codes could be viewed as different aspects of on an overall code/theme, but this shall be discussed further below. The one code shared across all three stances (neutral, Jose, and Miguel) was that new information is capable of creating a change in what is currently known:

> "I do not agree with them. Science like medicine is always changing. Nothing is ever set in stone because we learn something about our world and everything around us everyday."

> "Jose is correct because data is always coming in that changes things."

> "I agree with Miguel for the most part. When there is evidence behind a theory, it doesn't tend to change unless new evidence is found to prove otherwise."

Within these response types there was very little elaboration on the process by which this new information is believed to be acquired (or what was preventing this information from being discovered). Only a handful of responses (3 of 14, in this coding) discussed their beliefs in this regard, all of them attributing it to be knowledge acquired by the scientists directly testing the theory. This stance is likely the prevailing belief among these individuals, even when not explicitly stated. Interestingly, the three responses which did explicitly state how new information comes about agreed with Jose, for example:

> " I agree with Jose. Scientists are coming up with new ways to figure out how things happened, and why things are the way they are constantly."

It should not be of great surprise that another common code was the belief that nothing is set in stone in science (that there is always room for argument):

> "Jose and Miguel both have good points. Most theories are generally thought to be true. However, sometimes theories are disproven or added on to. Nothing is set in stone."

> “I agree with Jose but disagree with Miguel. Many theories and ideas are prone to change due to new technology like Jose says. I disagree with Miguel because everything can be argued and should have a counter argument.”

This response trait was unique to students who either were neutral, or agreed with Jose. A type of response trait which was predominantly favored by students who agreed with Jose was the belief that theories are readily changeable:

> “Jose is correct, Theories are theories, they are designed to change with new data. And for experiments already done and completed, there is no harm for argument, as it can help further prove it, or lay the groundwork for a new experiment.”

> “Jose, there’s no way of knowing 100% that a concept or theory is right, so it stands to think that the theories could change over time, multiple times.”

The basic views of these students wouldn’t be called unsophisticated, at least with respect to their views regarding theories/knowledge in science. So here is an example of students agreeing with an individual (Jose) whose stance is unsophisticated. Yet the arguments put forth by these students appear to favor a partially, if not entirely, sophisticated belief. Contrasting the above code is one in which students viewed scientific theories as anchors (difficult to change with little room for argument):

> “I agree partially with both. Hypotheses tend to fall in and out of favor and become paradigms; some are more hotly debated than others. Once a theory is established, it is usually less debated, though there are some theories which seem to be continually argued over.”

> “I agree with Miguel. Most of the theories and experiments that have been done have stayed the same over time or may have changed just a little bit. There isn’t much new information found on previous experiments or theories, just more research that proves it.”

As could be expected, this code type was seen mostly within students who agreed with Miguel, although several did appear to those who answered neutrally. It is important to mention that this stance is not one which reflects a belief of theories as unyielding facts (only one individual across all responses displayed this belief), just as theories being difficult to change. In regard to why these students believed theories difficult to change, it appears to be mostly based on the amount of evidence for a theory:

> “ I agree with Miguel. When a new theory is made prominent, it is not usually changed, just added to or modified.”

> " For the most part I agree with Miguel. I think that once they've tested something it usually is done and accepted. However, if more discoveries are made I believe that theories can be updated and improved upon."

> "I lean toward Miguel's side of the discussion. Usually if a theory has been formed, it takes significant evidence for scientists to change their minds."

It is of note that most of these individuals imply that there reaches a point where a theory is "formed" or "made prominent", this could be a topic of further investigating in the future. An aspect of student responses which was seen almost exclusively within individuals who agreed with Jose was the reliance on references to past/present scientific findings to justify science as being unstable:

> " I'd like to not agree with Jose, but I have to. I like to stay skeptical and keep an open mind, and when I look back on history and think about how incredibly confident intellectuals were in the flatness of our planet, I can't help but wonder what huge pieces of the puzzle we're missing."

> "I agree with Jose because science is very unpredictable, and I've seen throughout history science be wrong and other scientists having to reanalysis previous theories."

> I agree with Jose because yes, while tests and theories are proven the world around us is always changing so one cannot keep the same theory all the time. If one thing changes it can cause others to do the same and would then disprove that theory."

When this code did show up elsewhere, the students would balance the occurrence of perceived scientific instability in society with other considerations:

> "I don't know, it could go either way here, too. Science is a broad spectrum. I can understand the first student, in that what most affects us, like medical and medicine science, is often evolving. And that's the kind of science that most directly effects the life of people who aren't in the field."

In the student example above, they balance frequently changing science as being attributed to particular domains of science, and not science in general.

The overarching theme (and most prominent code) behind student responses to this question seems to be their views regarding science. Those who were neutral or agreed with Jose had a tendency to make claims such as:

> "I agree with both students to some extent, we should use scientific theory in all our studies of science, but we should also question these theories if we believe them to be false."

> “ I agree with Jose. Very commonly, especially with food studies it seems, scientist will seem to change their theories. It is always important to question theories, so that our scientific findings don’t become stagnant, because there will always be more for scientists to learn.”

or

> “ Again, I agree with both individuals to some extent. I agree with Jose because science is always changing. As advances are made to technology we are able to see things in a different light. But I don’t believe that scientist bounce back and forth. I agree with Miguel because scientist do continually test their theories to prove their validity. However, I do not believe that these theories are set in stone.”

> “ I agree with Jose because not everyone in settled on a theory, just like no one is settled in life. I believe that it is healthy to just back and forth to find different ideas that were missing from the last theories, etc. Because when you find an idea within a theory, you simply jump back to the theory that fit that idea and continue on. It is a never ending cycle.”

The former set of responses place a value on challenging/questioning the findings of science, while the latter set of above responses display a belief that scientific findings are being continuously tested and checked for accuracy. Views about science which were unique to students who agreed with Miguel was the idea that science “proves”:

> “Miguel because science is about proving theories to be what the world is about. All the different things that you can experiment on is permanent.”

> “Miguel, once a theory has been proven scientists don’t usually, if ever, back on those claims.”

> “I agree with Miguel because science is based on facts and observations not biased opinions.”

The manner in which the students here imply the idea of “proof” would seem to align with the view (unsophisticated) that scientific knowledge is set in stone.

All-in-all, however, responses in favor of either Miguel or Jose had a strong tendency to include both aspects which were sophisticated and unsophisticated, as it pertains to beliefs about science and knowledge presented within science.

EBAPS Question 30

The question as posed to students was:

> "Who do you agree with in the following discussion and why do you agree with them?
>
> Jessica and Mia are working on a homework assignment together . . .
>
> Jessica: Alright, we're done with problem 1. Let's move on to problem 2
>
> Mia: Hold on, I think we should try to figure out why the ball takes so long to reach the ground.
>
> Jessica: Mia, we know we've got the right answer, it's in the back of the book. If we didn't understand the problem we wouldn't have gotten the right answer.
>
> Mia: No, I think it's possible to get the right answer without really understanding what the answer means."

This question is intended to probe student epistemologies in the same manner as does EBAPS question 30.

There is one predominant issue within the problem, as seen in the written student responses. This concern is that of time, specifically, in how it plays into the role of reflecting upon the problem:

> "It's true one can not understand but have the means of achieving the correct answer. However, in the terms of homework, they need to move on and can look into it at greater lengths later."

As can be gleaned from in the example above, several students consider whether or not it is worth investing the time to work the problem in further detail. However, it is likely that this is not a reflection of their epistemic sophistication so much as it is simply time management in a busy student schedule. Students who respond in this manner do have a strong tendency to acknowledge the value of reflecting upon the problem. As such, this problem has a rather concerning implicit factor (time management) beyond that of the epistemological domain, which may skew the exploration of these student beliefs. Having expressed this concern, it is of note that majority of the students responding in this manner did still have a sophisticated response as they would declare agreement with Mia.

There were 52 total written student responses to this question. Out of these responses, 2 were neutral (agree/disagree with Mia and Jessica), 5 agreed with Jessica (unsophisticated), and 45 agreed with Mia (sophisticated). The first aspect to student responses that shall be discussed has already been outline above, that being the focus on time:

> "Jessica is right. If they have time at the end, they should revisit the topic and try to understand it together, but it is more important to get the work done right away, then go back and fix the rest later."

> “Mia is correct to gain a full understanding of the material. However, if they are working in class or under some kind of time constraint it may be more efficient to move on to a different problem for the time being.”

The pressure that students feel due to the demands of college cannot be denied given the nature of this type of response trait. Students alter their behavior because they need to manage their time accordingly. If students believe they can obtain a satisfactory grade by only being able to get the right answer, it is likely that they will do so. However, most students who are hinting at being pressured by time still acknowledge the wisdom of reflecting back upon the problem. Fundamentally, there was only one response involving time as a factor which portrayed an unsophisticated belief regarding knowledge:

> “I agree with Jessica. If you have gotten the right answer and checked it to make sure it is correct, then you must have used the right process. There is no need to waste time, therefore you can move on to the next question.”

Another common occurrence among student responses was that of a future-problem philosophy:

> “I agree with Mia because you should always strive to understand fully or else you won’t know it come test time.”

> “I agree with Mia because taking the easier route will bite you in the butt soon before you know it. When you take the long route within the problem, it will be easier to solve in the future because you took steps to find the answer rather than flip to the back of a book.”

Here, students place importance on understanding the intricacies of the problem now because it will be of benefit in the future. Typically, this importance is usually justified with the belief that there will be a similar problem that arises either in a later homework or an exam. Sophistication beyond this was not seen within this code. Ideally, students might justify the need for understanding not only for future class work, but for work outside the current classroom. There was a notable amount of difficulty in the identification of patterns within student responses. Due to this, there became more emphasis on words within student responses as well as the overall attitude of the response. The following two codes were acquired with this general approach. The first aspect of student responses that will be discussed are those which put an emphasis on concept and “why”:

> “I agree with both of the students. At times I just want to move onto the next problem when I know that we have the right answer but it is also important to take time to really understand the whole concept of the problem and make sure that you understand the problem and everything that is involved in the problem.”

> “I agree with Mia because it is important to understand what it is that you are doing and why you are doing it, not simply coming to the correct conclusion.”

> “Mia. The math can accidentally just workout at times but concepts should be fully understood before moving forward. Math and sciences build upon themselves as concepts become more advance so a good amount of understanding should be attained at each level.”

This was the most frequently occurring code within the written responses and gave more thought regarding why it is important to achieve understanding. Usually this thought manifested in sophisticated views about the value of concepts and discovering reason. Although the above category may be vague, it is made clearer by the presence of a complementary code. The second of the two aforementioned codes has students placing an emphasis on “how”, as opposed to “why”:

> “I agree with Mia. I have gotten the right answer without understanding the steps to get there.”

> “I agree with Mia. It’s better to understand how you got your answer so you can do it again in the future.”

As seen in the examples above, there seems to be a focus on the process by which the answer is acquired, which is dissimilar to an attention around concept and reason as previously discussed. Again, this dichotomy of “why” versus “how” is something derived by the researcher and are traits not commonly explicit within student responses. For example, when the student above states “[. . .] how you got your answer so you can do it again [. . .]” would seemingly imply a view for which the student is desiring obtaining knowledge of the algorithm by which the problem is solved. It is not without possibility, however, that the student is indeed referring to more sophisticated constructs such as concepts when making the statement. Another focus of students within their responses was one which seemed to reflect aspects of metacognition. These aspects being a discussion of what it means to know:

> “I agree with Mia. You may know the answer because you memorized and regurgitated it but not understand it. Someone who understands the answer to a question is able to explain it in more ways than one.”

And a general reflection on the problem and their reasoning:

> “I agree with Mia. Speaking from experience, I can say that it is most definitely possible to answer a question right without fully understanding the problem. This can happen due to lucky guesses or flawed understanding. Many times I have correctly answered a question without knowing how exactly that answer is correct. When working through the problem,

> I came up with a reason to justify how the answer was right, only to later find out that my reasoning with completely wrong, even though the answer was right."

When it pertains to an overall theme within the student responses to this question, it is tempting to categorize them as varying degrees of problem solving sophistication. Specifically, the metacognitive sophistication as it pertains to problem solving, i.e. is it beneficial to reflect upon the problem once the right answer is achieved? However, there seems to be a more general trend to these responses than just metacognition. Overall, these responses seem to be student beliefs regarding how to handle potentially conflicting claims (right answer vs. correct understanding) and uncertainty in knowledge. Within the students who answer unsophisticatedly, there seem to be two trends, that of either a heavy emphasis on time management (as previous discussed) and/or an emphasis on interest. That is, if the student is interested in the material, then they will reflect upon whether they truly understood it.