Spring 5-2018

Echoes of the Past: The Effect of Background Experience on Far Transfer

Graham H. Hummel-Hall
University of Maine, dreyfus2006@hotmail.com

ECHOES OF THE PAST: THE EFFECT OF BACKGROUND EXPERIENCE ON FAR TRANSFER

by

Graham Hanson Hummel-Hall

B.S., University of Wisconsin-Madison, 2015

A THESIS

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science in Teaching

The Graduate School

University of Maine

May 2018

Advisory Committee:

Christopher Gerbi, Associate Professor of Earth and Climate Sciences, Advisor

Craig Mason, Professor of Education and Applied Quantitative Methods

John Thompson, Professor of Physics and Cooperating Professor of STEM Education

ECHOES OF THE PAST: THE EFFECT OF BACKGROUND EXPERIENCE ON FAR TRANSFER

By Graham Hummel-Hall

Thesis Advisor: Dr. Christopher Gerbi

An Abstract of the Thesis Presented
in Partial Fulfillment of the Requirements for the
Degree of Master of Science in Teaching
May 2018

Far transfer is the application of knowledge learned in one setting to a problem in a very different setting. This multi-method study looked at far transfer in humans and whether it could be facilitated, inhibited, or remain unaffected by the number of courses or years a student at a university spent learning about the subject matter of the knowledge being transferred. Through quantitative and qualitative analysis of pretest and post-test data from an introductory undergraduate earth science course, I found that students with more physical science background experience more frequently engaged in successful and accurate transfer of physics information to novel questions relating to plate tectonics. I also found evidence that a high amount of previous physical science experience seemed to promote transfer later in the earth science course. However, due to the sample size of the analyzed student responses, I believe that my results are preliminary and I encourage more research to be done on the topic with larger sample sizes.

ACKNOWLEDGEMENTS AND DEDICATIONS

First and foremost, I would like to thank my thesis committee members Christopher Gerbi, Craig Mason, and John Thompson for their support, input, and flexibility. It meant so much to me to have a team of superiors rooting me on and supporting me every step of the way, and their advice was invaluable. I felt in awe of them every time I was in a room together with the three of them, and I hope one day that I might be to my students what they were to me: role models, teachers, sources of guidance, and friends. I would especially like to thank Chris, my advisor, for all the patience he has shown me and all the guidance he has provided me with. He has been encouraging me ever since I first came to him with a thesis idea, and this thesis, as well as my own personal growth throughout the thesis process, could not have been possible without him.

I would also like to thank Scott Johnson for his assistance and cooperation with Chris in data collection. Without his help, there would have been a lot more work and a lot more time for me to tackle in this study. Chris and Scott were responsible for amassing the pretest and post-test data that I based my entire study on, so I owe to them a lot of gratitude and appreciation.

This thesis would not be what it is without the encouragement and support of my friends and family, but I would especially like to thank and dedicate this thesis to my mother, Cynthia Hummel, who rooted for me and supported me every step of my life's journey. It was through her that I first learned of theses, and in writing mine I did my best to create a thesis document that may have been something like what she may have wrote all those years ago when she wrote hers. It is a great tragedy that she did not live two more years to see my academic career finally come to an end. I hope this thesis would make her proud.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

T: "Temperature (concept category)"; used in the Short-Answer analysis

ρ (rho): "Density (concept category)"; used in the Short-Answer analysis

T-*ρ*: "(Relationship between) Temperature and density (concept category)"; used in the Short-Answer analysis

V: "Vertical motion (concept category)"; used in the Short-Answer analysis

LIST OF DEFINITIONS

College Experience: The total number of physical science courses that a student previously took in undergraduate school (i.e. physics courses, chemistry courses, etc.).

Course-Year(s): Imaginary unit of experience invented by the author to equivalate the amount of content learned in a semester undergraduate course (counted in number of courses) to the amount of content learned in a yearlong class in secondary school (counted in years); under this system, when determining one's amount of experience, 1 course + 1 year = 2 course-years.

Demographics Questions: Questions 22 and 23 of the pretest and Questions 17 and 18 of the Post-Exam III survey. While there were several questions on the pretest and the Post-Exam III survey that asked about demographic information, *demographics questions* in this study refers specifically to the aforementioned four questions as they were the only questions about demographics used for this study. Questions 22 from the pretest and 17 from the Post-Exam III survey asked about previous physical science experience in college, and Questions 23 from the pretest and 18 from the Post-Exam III survey asked about previous physical science experience in middle school and high school.

Experience Group: One of two bins that students were placed in based on their reported amount of previous experience in physical science. The *low experience group* had two or fewer courses/years/course-years of previous physical science experience, whereas the *high experience group* had more than two given units of physical science background.

Post-Exam X *Survey*: One of three post-test surveys administered to students following each corresponding exam in the introductory earth science course. The first two exclusively included questions about earth science, whereas the third also included questions about background experience, identified gender, etc. that were previously asked on the pretest. Of interest to this study are the "Post-Exam I survey," which included all of the post-test questions involving plate tectonics investigated by this study, and the "Post-Exam III survey," from which the physical science experience of students responding to the Post-Exam I survey was gleaned. The "Post-Exam II survey" does not appear in this study beyond a brief mention.

Secondary School Experience: The total number of academic years that a student during which a student was taking at least one physical science class (i.e. Physical Science, Physics, Chemistry, etc.).

Total Experience: The sum of a student's background experience in the physical sciences in middle school, high school, and college. In this study, total experience is reported in units of *course-years* due to secondary school experience and undergraduate experience being reported in different kinds of units.

1. INTRODUCTION

One of the most important goals for formal education is that students will apply what they learn in the classroom to other settings, a phenomenon known as *transfer*, sometimes called *transfer of learning* or *transfer of training*. The study of transfer in humans is thus crucial for both improving education as well as for confirming that graduating students under current educational practices do indeed engage in transfer. Given its importance as both a constructive study and a reflective study, a lot of research has been done on transfer over the years on a number of different topics (e.g., Barnett & Ceci, 2002). One area where transfer research is particularly helpful is cross-cutting relationships between fields (e.g., Bassok & Holyoak, 1989). It is in that area that I will now pose a question and subsequently spend the rest of my time helping to scientifically answer. If two fields are related, is taking a lot of classes in one field beneficial for thinking about the other? For example, can learning about physics prepare one to think about geoscience, perhaps about a novel geophysics scenario? Can learning about physics make it easier for one to learn about geoscience?

This Study

The question posed by this study is thus: "How does the amount of background experience one has in physical science affect their performance on questions about earth science?" If the amount of background experience one has has some effect on the way one thinks about questions about earth science in undergraduate school, I posit that the performance in an introductory earth science course when answering a question about plate tectonics will be different between undergraduate students with significant amounts of physical science experience and those with only a little physical science experience. Understanding the impact of the amount of experience on transfer to a novel but related subject has many utilities in improving education in the United

States and abroad. For example, it may inform colleges whether they should recommend students take physics before taking any earth science courses, and how much physics and chemistry they should take. Ideally, educators would want students to transfer information between related topics (as so heavily emphasized by the "cross-cutting relationships" outlined in NGSS[1] K-12 standards in the United States), so it is important to know whether students are doing that in introductory college courses or if something needs to change in order to better facilitate that transfer.

Transfer can manifest itself in different degrees. I was interested in the frequency of *successful and accurate transfer*. *Successful transfer* I will define as an event in which the brain of an animal (in this case *Homo sapiens*) notices cues in the environment, searches its memory banks for relevant previously learned information, and successfully recalls that information, regardless of whether that information is actually accurate to reality or not. For example, when asked about how hot air balloons function, a student who hears the word *hot* and applies what they learned about heat conduction would have engaged in successful transfer, even though hot air balloons actually function according to heat convection. *Accurate transfer* I will define as transfer in which the transferred information is "correct" (or accurate to what society currently understands about the world). The previous example was an example of inaccurate transfer. Had the individual recalled information about heat convection rather than heat conduction, that would have been more accurate.

Using a mixture of quantitative and qualitative approaches, I found that students in an introductory earth science course did engage in successful and accurate transfer more frequently

[1] "Next-Generation Science Standards"

when they had a lot of previous background in physics than when they did not, but also that their amount of experience was an influencing variable with (in some cases) only a very small effect. These results imply that more attention should be given by students to quality of instruction rather than quantity and that more could be done by physical science and earth science instructors to encourage far transfer.

2. BACKGROUND

Far Transfer

For such a phenomenon as transfer to be possible, a student's brain would have to take information learned in one setting—the physics classroom—and be able to recognize cues to apply that physics knowledge in another setting—perhaps a geoscience classroom—and then recall that information as it was originally learned so that it may be applied. In other words, that student would have to transfer their knowledge to the new setting. Whether such a task is possible for a nervous system has only been studied in a choice few taxa, all of which are mammals (Phylum Chordata, Class Mammalia). However, it seems to be the case that transfer can at least be performed by the euarchontoglires *Mus musculus* (house mice) (Kurt & Ehret, 2010) and *Homo sapiens* (humans) (i.e., Fuchs et al., 2003; but see the following paragraph). The study of transfer in non-human mammals such as *M. musculus* is still in its infancy. For the rest of this paper, I will be focusing on research on transfer in *H. sapiens* and the findings thereof.

Over the past few decades, there has been controversy over whether humans truly engage in transfer (as summarized by Barnett & Ceci, 2002). At a very broad level, some argue that the frequency of transfer is incredibly rare in humans and comparable to "volcanic eruptions and large earthquakes" if not rarer, as evidenced by the ratio of the number of professors (whose job it is to transfer previously learned material to new scenarios) at universities to the number of professors who have successfully made novel discoveries (Detterman, 1993). As an example of evidence suggesting humans struggle greatly with transfer, it was found in 1974 that college students who had previously learned to solve either the classic missionary-cannibals problem or an identical jealous husbands-wives problem struggled or were completely unable to solve the other problem even though the two problems were nearly exactly the same (Reed, Ernst, & Banerji, 1974). Brown, Kane, & Long (1989) argue that at least some difficulties with observing transfer in humans is because of (as they call it) the common "And Now for Something Completely Different" paradigm. This paradigm involves a researcher presenting a subject with two related concepts but withholding any and all hints as to how the two concepts are connected, expecting that the subject will recognize the connection entirely on his or her own. By instead presenting problems in a manner similar to how they are presented in formal educational settings—such as a progression from a simple problem to a problem explicitly known to be an analogy to the first problem, then back to the first problem, and then finally to a novel but related problem not explicitly stated to be related to the first two—transfer can be much more easily observed in humans (Brown, Kane, & Long, 1989).

Studies on *H. sapiens* that have successfully exhibited transfer have revealed a wide variety of different categories of transfer. For example, a distinction can be made between *positive transfer* and *negative transfer*, where background experience enhances or hinders one's ability to learn something new respectively (Royer, 1979). In the case of *forward-reaching transfer*, an individual may learn something, think about how that knowledge may be applied to other situations, and then later transfer that knowledge; in the case of *backwards-reaching transfer*, an individual may encounter a novel scenario and think back in order to identify what concepts he or she may have previously learned that might be applicable before transferring them (Salomon & Perkins, 1989). Transfer may be *horizontal* if an individual must apply a concept to a problem of similar complexity, or it may be *vertical* if the new problem is of a different level of complexity compared to the source (Barnett & Ceci, 2002). Transfer may be *specific*, in which the particular contents of a lesson are transferred by the learner (e.g., the contents of a list of countries), or it may be *non-specific* or *general* if the learner instead transfers the general skills they practiced while conducting the lesson (e.g., the ability to memorize the contents of a list, hold one's attention on a list, etc.) (Detterman, 1993). An individual engages in *near transfer* when they transfer their knowledge to a setting that is similar or identical to the one in which that knowledge was originally learned. An example of near transfer would be when a student applies information they learned in October in a high school physics class to a November test about that material in the same class. By contrast, *far transfer* involves transferring material learned in one setting to another setting that the individual perceives as very different from the original setting. An individual would be engaging in far transfer if they applied that knowledge learned in October in Physics to something in their English class, or perhaps to a round of tennis with their friends (Royer, 1979). Often when transfer is discussed, it is far and/or general transfer of deep structures that is of interest, rather than near and/or specific transfer of surface-level structures (Detterman, 1993). It is worthwhile thus to consider that the boundary between what qualifies as near transfer and what qualifies as far transfer is subjective and poorly defined as a whole by the scientific community. For the purposes of my study, *far transfer* will be henceforth treated as "transfer that occurs in an environment that the actor thinks is completely different from where the material was originally learned"—for example, a student may perceive a math class as a completely different setting from an English class because they believe "math is not done in an English class," and a kitchen as even more different.

Barnett & Ceci (2002) constructed a taxonomic system to help break far transfer down into more-distinguishable pieces. They utilized two classification systems: one based on the content that was being transferred, and one based on the contexts in which the content was learned and in which the content was being applied. Far transfer, they argue, can be broken down by how generalized the content is, what improvement is hoped to be attained through transfer (e.g., "Will the individual perform an action faster if transfer is successful?", "Will the individual be more accurate if transfer is successful?", etc.), and whether the transfer requires the individual to recall past content or to recognize that content. Similarity between the context in which something is learned and the context in which it can be applied can be measured by how similar the knowledge domains are (e.g., biology and chemistry versus biology and music history), how similar the physical contexts are either at the micro scale (e.g., the same classroom versus two different classrooms) or the macro scale (e.g., a classroom versus a beach), how similar the temporal contexts are (e.g., a single day apart versus multiple years apart or two time-constrained contexts versus a time-constrained context and a context that is not time-constrained), how similar the functional contexts are (e.g., both are academic versus one is academic whereas the other is for obtaining sustenance from frozen food), how similar the social contexts are (e.g., individual is alone in both contexts versus individual learns content alone but is asked to apply it while in a group), and how similar the behavioral contexts are (e.g., both are oral versus individual learns content in a lecture but has to apply it through wood carving) (Barnett & Ceci, 2002).

Because research on transfer has still only recently extended beyond *H. sapiens* (Kurt & Ehret, 2010), ideas for the mechanisms of far transfer are still largely restricted to anthropocentric theories of learning. Notably, the mechanisms that enable far transfer have been linked to schema theory (Gick & Holyoak, 1983) and information processing theory (Cooper & Sweller, 1987). It seems that humans require awareness of relationships between two problems in order to transfer knowledge applicable to one problem in order to solve the other problem. Even when they are aware that two problems are conceptually related, humans still do not consistently engage in far transfer of knowledge unless they have well-constructed schemata of what that knowledge can be applied to (Gick & Holyoak, 1983) and have already sufficiently abstracted the concept in question (Salomon & Perkins, 1989). Fuchs et al. (2003) nicely summarizes the relation of far transfer to schema theory as the need for humans to “develop categories for sorting problems that require similar solutions” before they are able to engage in far transfer. Cognitive load in one’s working memory also seems to affect human far transfer. Streamlining problem-solving by automating one’s understanding of the rules and limitations of a problem was found to open up space in the brain’s working memory for planning and to facilitate transfer (Kotovsky, Hayes, & Simon, 1985). Alternatively, increasing cognitive load hinders the ability for human children to engage in transfer (Karbach & Kray, 2009).

Cooper & Sweller (1987) and Fuchs et al. (2003) provide very nice overviews and literature examples of these connections to anthropocentric theories of learning. I will thus instead turn to how such mechanisms in humans may be nurtured or supported so that educators may increase the frequency at which their students engage in far transfer. While several studies show humans exhibiting far transfer, there are also several studies about transfer in which transfer does not occur (Barnett & Ceci, 2002). If engaging in far transfer is so variable in humans, how can we help them to apply their previous knowledge to novel situations?

Does the source of knowledge affect its transferability? In 1989, Bassok and Holyoak presented high school and college students who had background experience in arithmetic-progression algebra and those who had background experience in thinking about one-dimensional motion when acceleration is constant with word problems about either domain. What they found was that where a student learns about a concept affects whether it is internalized as a generalized concept or as a specialized concept, and that concepts are more transferrable when generalized. Math classes, they found, better supported transfer to physical science problems than physics did to math problems. This is believed to be because math classes feature a wider variety of problems that encourage students to apply what they have learned to a great variety of scenarios, whereas the scenarios presented to students in Physics are typically more narrow in scope (Bassok & Holyoak, 1989).

Does age affect whether *H. sapiens* can engage in transfer? The question is an important one, particularly for schools at the K-12 level and lower. Karbach and Kray in 2009 investigated in Germany whether the ability to engage in non-specific transfer, in this case of executive control of the human nervous system, varied by age. Their sample was sexually heterogeneous and included 56 children (ages 8-10 years old), 56 young adults (18-26 years old), and 56 aged adults (62-77 years old). These subjects were given various executive tasks—for example, in one task they had to identify whether they were looking at a picture of a fruit or a vegetable; in another, they had to identify whether a picture was small or large; in a third, they had to indicate how many alphabetical characters were being presented at once (a *Number-Stroop test*); and so on and so forth. They were then given training in different but similar tasks, and presented with the original tasks again to see if the subjects would transfer how they learned to control their executive function to the original tasks. They found that while the two groups had deficits in their ability to select and maintain sets of tasks compared to young adults, children and aged adults still engaged in successful near transfer of training that involved switching tasks every second task. Being required to engage in variable training tasks promoted near transfer in young and aged adults but hindered it in children. However, age was found to have no effect on far transfer of task-switching training. This suggests that the mechanisms of far transfer are not age-dependent, and so while attention should be given to age when considering near transfer in *H.*

sapiens, such attention is not necessary when considering far transfer. As an interesting aside, this study also found that verbal self-instruction did not promote near transfer (Karbach & Kray, 2009).

How does the depth at which a student understands a topic impact the frequency of far transfer? Transfer seems to be promoted much more in young *H. sapiens* when they are asked to explain why they have identified something as fitting a general theme compared to when an instructor explains how that thing or concept fits the theme (Brown & Kane, 1988). Brown & Kane (1988) and Brown, Kane, & Long (1989) together have shown that at least in children, it seems that deeper understanding of a concept facilitates far transfer of knowledge about that concept, as well as some prior understanding of the context in which the concept is being applied (but see Barnett & Ceci, 2002, which suggests that these two studies exhibit *near* transfer rather than *far* transfer); consistent with Karbach and Kray's 2009 findings, the latter conclusion has also been found to be true for adults (Schliemann & Magalhães, 1990).

Earth Science and Physical Science

The earth sciences and the physical sciences, while both fields of science widely included among others underneath the STEAM[2] umbrella, have very different histories.

The physical sciences encompass physics and chemistry, largely two sides of the same coin. Physics, the "original" science as far as Western cultures are concerned (although it would be centuries before true scientific approaches would be applied to it), started out as the study of more-or-less anything in the natural world, one of the three branches of philosophical thought

[2] *STEAM*, or "Science Technology Engineering Art Mathematics," is a grouping used in the United States when talking collectively about those fields in the context of education or careers.

delineated by the Cypriot Zeno more than 2,200 years ago. Modern physics today has been reduced to those aspects of the natural world that are not the focus of other major branches of science: things like motion, electricity, waves, magnetism, thermodynamics, and time. Chemistry is much more recent, although aspects of it have been practiced since the days of the early Egyptians. It is a study largely focused on atoms, molecules, and compounds and their interactions. As modern physics also still includes many topics directly relevant to compounds and their components, physics and chemistry share many characteristics and are often grouped together as *physical science(s)*.

The earth sciences encompass many different fields of study united by their focus on abiotic (and largely macroscopic) aspects of nature. Geoscience, meteorology and climatology, and astronomy are among those fields considered to be *earth science(s)*. These fields have very different origins. Across many cultures, astronomy, the study of parts of the universe beyond Earth's atmosphere, was one of the earliest sciences. In Western cultures, geoscience has its origins tied deeply to naturalism (the progenitor of modern biology) and can trace its history to the 18th Century. It can sort of be thought of as the study of the abiotic Earth, but can be applied to other planets and often crosses over into biotic subject matter when the focus of study is Earth's history. The most common way to describe geoscience (or more accurately *geology*, as it used to be called) is as the study of rocks and minerals, but this description is often frowned on by geoscientists for being too limited. Together, earth scientists might study plate tectonics, volcanoes, the Earth's atmosphere, the chemical composition of Jupiter, ocean currents, or the origin of limestone, among other things.

The Earth is composed of a wide variety of chemicals and is a part of the natural world. It stands to reason, then, that despite their different histories, the earth sciences are fundamentally connected to the physical sciences. *Geophysics* is a cross-cutting field that considers the physics of abiotic Earth materials. For example, a geophysicist may be interested in the driving forces behind plate tectonics. Some planets, such as the Earth and historically Mars (Sautter, et al., 2015), have surfaces comprised of solid plates of rock moving and being created and recycled atop the more slushy rocks underneath, a process known as *plate tectonics*. There is still a lot to be learned about the drivers of plate tectonics, but the current working theory is that the plates are part of a sort of underground conveyor belt extending from the planet's surface through the mantle. To use the Earth as an example, the Earth's core heats up the overlying mantle material. The molecules inside the heating material get more excited and bounce off of each other more forcefully, causing the material to expand and each cubic centimeter of the material to be thereby less dense. The surface of the Earth is much cooler—and thus denser—than the mantle material underneath it. Dense material sinks underneath less dense material (for example, an iron ball will sink to the bottom of a bucket of water, which is less dense than iron, whereas an even less dense wooden ball will float on top of the water), so mantle material close to the Earth's crust will sink down towards the core, displacing the warmer material below and pushing it upwards. As the warmer material rises, it gets farther away from the core and cools down, until eventually it is so dense that it too begins to sink and push other material up. The cycling materials nudge along the brittle pieces of the mantle and crust—the plates—above them. Physicists call this cycle of warm, less-dense material rising, cooling down and becoming more dense, sinking, and warming up and becoming less dense *convection*. Mantle convection is just one of endless examples of how earth sciences connect to physical sciences.

At the turn of the 20th century, American schools widely treated earth sciences (as well as life sciences) as courses unrelated to physical sciences and more akin to nature appreciation courses. Today, it is firmly understood that the physical sciences provide the ground that all other sciences are based on, and many American colleges and universities today require students pursuing earth science degrees to spend some semesters studying the physical sciences. While Physics and Chemistry are required for these students, the historical divide between how earth sciences and physical sciences are perceived largely remain among students. In addition, American college courses are sometimes taught very differently from K-12 classes. This means that far transfer may be required in order for a student to make use of their physical science knowledge when thinking about earth science, especially in the case of physical science knowledge acquired in secondary school.

3. METHODS AND RESULTS

Data Collection

The data used for this study came from a pretest/post-test survey presented to Earth Science 103 students at the University of Maine in 2014 about the material covered by the course. I was not involved in the data collection process. The survey was administered in four stages: a pretest and three post-tests that were administered over the course of the semester as students learned the relevant material. These post-tests are referred to as the *Post-Exam I survey*, the *Post-Exam II survey*, and the *Post-Exam III survey* respectively, according to which class exam the post-test was administered after. Each survey included both multiple-choice and short-answer questions, although the actual number of questions varied between each survey.

For Questions 2 through 5 on the pretest, students were given a diagram showing a cross-section of one quadrant of the Earth on which a divergent boundary and a subduction zone were clearly visible (see Figure 1). Question 2 asked students to pick which arrow from a selection of four choices best represented the direction of the mantle's motion at a subduction zone boundary (labeled as *X* in Figure 1). Question 4 was similar but asked about the direction of the mantle's motion at a divergent boundary (labeled as *Y* in Figure 1). Students were hoped to respond "D" (motion towards the planet core) for Question 2 and "A" for Question 4 (motion away from the planet core). Questions 3 and 5, which were not analyzed in my particular study, asked students to explain their answers to Questions 2 and 4 respectively. Question 6 on the pretest departed from the diagram and asked students to provide a short-answer response as to how temperature and density differences drive terrestrial plate tectonics. The ideal response was for students to identify that the material in the deep part of the Earth's mantle near the core heats up, causing it to become less dense and rise up towards the Earth's surface, cooling down and becoming denser in the process until eventually the material sinks back down towards the core, creating a convection cycle. Questions 20 through 23 on the pretest asked students to quantify their background experience in earth science (Questions 20 and 21) and physical science (Questions 22 and 23) in college and in middle school and high school respectively. How students could respond to these questions will be touched on later during the discussion about how student background experience was calculated.

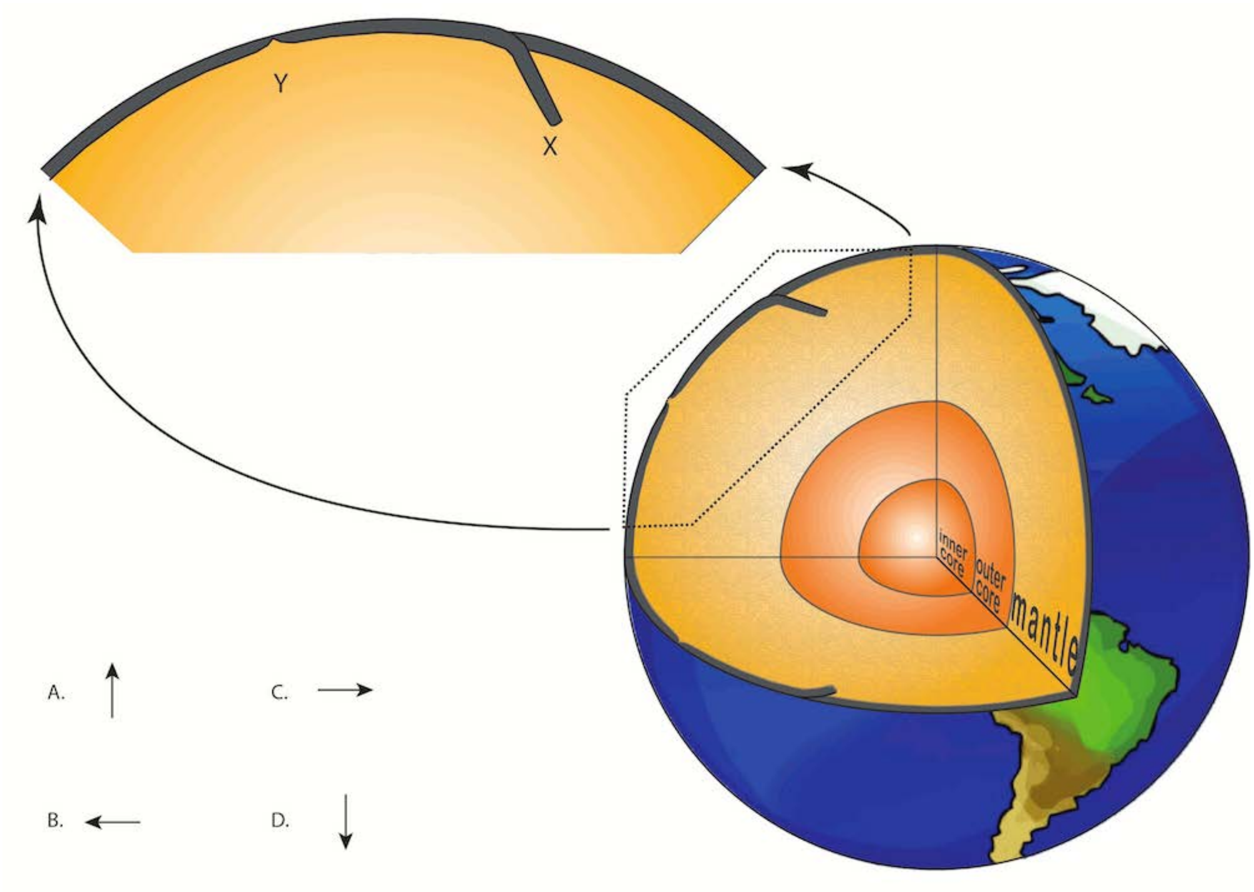

Figure 1. Image associated with Questions 2 and 4.

The questions on the pretest were each later asked again on one of the three post-tests. Questions 2 and 4 on the pretest reappeared on the Post-Exam I survey with the same correct answers as they had on the pretest. Question 6 on the pretest also reappeared on the Post-Exam I survey as Question 7. Questions 20 through 23 on the pretest reappeared as Questions 15 through 18 on the Post-Exam III survey. When they took the pretest/post-tests, students were attached with an anonymous numerical identifier so that a student's responses on any of the post-tests can be paired to the same student's responses on the pretest.

Calculation of Experience

The first step in the analysis of the multiple-choice responses was to place students into bins based on their physical science background. Students self-reported their physical science background in Questions 22 and 23 on the pretest and again in Questions 17 and 18 on the Post-Exam III survey. These questions will henceforth be referred to in this study as the *demographics questions* (although other questions about demographics—i.e. identified gender, whether the course was taken live or online, etc.—were asked on the survey, they are not of known relevance to this study and will not be discussed further). On the survey, students could respond to the demographics questions with discrete numbers ranging from 0.0 to 3.0; if they wanted to respond with a number greater than 3.0 they could select “More than 3.”

As I mentioned in the Introduction, the differences between secondary school and college may enhance the perceived differences between where one learned about physical science and the setting in which students responded to earth science questions. I thus considered three different approaches to measuring background experience: measuring just a student’s physical science experience in secondary school (*secondary school experience*), a student’s physical science experience in college (*college experience*), and a student’s total physical science experience between the two (*total experience*).

A notable issue for calculating total physical science experience was that the questions about college experience (Question 22 on the pretest and Question 17 on the Post-Exam III survey) asked students how many *courses* they had taken in the relevant subject (in intervals of 1.0), whereas the questions about middle school and high school experience (Question 23 on the pretest and Question 18 on the Post-Exam III survey) asked students how many *years of study* they had done in the relevant subject (in intervals of 0.5). This meant that the measured units for college experience (number of courses taken in the relevant subject) were different from the measured units for secondary school experience (number of years of relevant instruction). In order to combine the two into a single measurement of a student's total experience in the relevant subject, one college course was equated to one year of secondary school experience. This was done because the amount of content covered in a single semester-long college course is roughly the same as the amount of content covered in a single year-long middle school or high school course. Because the two units can be approximated to have a 1:1 relationship, they will be treated as functionally synonymous and will be referred to henceforth as *course-years*.

Another problem for calculating a student's total experience is that not all the possible responses to the demographics questions were on an interval-ratio scale: the choice "More than 3" does not carry any information about a student's experience apart from that the number of courses or number of years totaled more than 3 courses or years respectively. Because it was impossible to deduce, for example, whether a particular student responding "More than 3" had taken 4 courses in college or 5, any student's total experience greater than 3 course-years was treated as simply ">3 course-years."

After making the necessary adjustments, I looked at student physical science experience right at the start of Earth Science 103 in three ways: college experience (based on responses to

Question 22 on the pretest), secondary school experience (based on responses to Question 23 on the pretest), and total experience (based on the sum of a student's college and secondary school experience). Initially, my plan was to split each statistical population up into three bins; however, while over 114 students took the pretest and 79 students took the Post-Exam III survey, there were too few students remaining after removing those that did not respond to the demographics questions to justify splitting the students into three bins (I will elaborate on the topic of sample size in Chapter 4). Ultimately, I placed students based on their reported background experience into one of two bins: those with experience (in the corresponding units) of 0.0 to 2.0 and those with experience greater than 2.0. I will henceforth refer to the group of students within each bin as *experience groups*. So, for example, when looking at student responses based on their secondary school physical science experience, students were split up into those who took 0.0 to 2.0 years of physical science courses in middle school and high school and those who took more than 2.0 years—making a total of two experience groups. The 0.0-2.0 bin represents the experience group consisting of *students with little to no (college/secondary school/total) physical science experience*, whereas the >2 bin represents *students with some or a lot of physical science experience*.

The physical science experience reported by students on the Post-Exam III survey was calculated and organized in the same way as that reported on the pretest. The inconsistency between responses to the demographics questions on the pretest and the responses to the same questions on the Post-Exam III survey in both reported experience and in the number of total responses means that Post-Exam III survey experience groups should not be compared to pretest experience groups (see Table 1). Only 48 (about 47%) of the 103 students who reported their past physical science experience on the pretest reported their past physical science experience on the Post-Exam III survey. This is partially due to participation levels—of the 114 total students who took the pretest, only 55 of them (roughly 48.25% of the students who took the pretest) took the Post-Exam III survey. In addition, 3 students who answered the demographics questions on the pretest *did* take the Post-Exam III survey but did *not* answer the demographics questions. These students were thus only included in the data analyses that used the responses to the pretest demographics questions to determine physical science experience.

	Consistent	*Inconsistent*	*Lost*	*Gained*
Secondary School Experience	17	31	55	5
College Experience	39	9	55	5
Total Experience	20	28	55	5

Table 1. Consistency between responses to the demographics questions on the pretest and on the Post-Exam III survey by students who took the pretest. The "Secondary School Experience" row compares responses to Question 23 on the pretest with those to Question 18 on the Post-Exam III survey; the "College Experience" row compares responses to Question 22 on the pretest with those to Question 17 on the Post-Exam III survey; the "Total Experience" row compares the sums of the two pretest questions with those of the Post-Exam III survey questions. A student's responses were "inconsistent" if their response on the pretest differed from their response on the Post-Exam III survey, "lost" if they only responded on the pretest but not the Post-Exam III survey, and "gained" if they only responded on the Post-Exam III survey but not the pretest. Students who did not respond to the demographics questions on the pretest nor those on the Post-Exam III survey are not included.

While students answering the demographics questions on the pretest but not the Post-Exam III survey were the largest source of the inconsistency between responses to the demographics questions on the two surveys, there were two other sources. Five students took the pretest but only reported their past physical science experience on the Post-Exam III survey. They were thus included in the data analyses that only used the responses to the Post-Exam III survey to determine physical science experience. Finally, some students changed their answers to the demographics questions between the pretest and the Post-Exam III survey, mostly for those questions involving their secondary school experience. While these students were included in all the data analyses, it is important to note that they may have been included in different experience groups depending on if an analysis used the responses to the pretest demographics questions or the responses to the Post-Exam III survey demographics questions in to determine past physical science experience.

Some aspects of Chapter 4 hinge on there being no negative correlation between secondary school experience and college experience. I performed a τ test on the responses to the demographics questions to check this (Kendall, 1938). This study posits that if there is a dependency between secondary school experience and college experience, statistical analysis should produce a Kendall's rank correlation co-efficient that is not equal to zero (mathematically, $\tau_b \neq 0$). At a significance level of 0.05, I found that the Kendall's rank correlation co-efficient for both the pretest ($\tau_b = 0.15$, $p = 0.177$) and the Post-Exam III survey ($\tau_b = 0.09$, $p = 0.429$) were not statistically significantly different from 0. This means that it cannot be ruled out that reported secondary school experience is independent of reported college experience.

Multiple-Choice Data

Analyzing Multiple-Choice Data

To analyze the pretest multiple-choice data, I first constructed a raw frequency table of responses to Questions 2 and 4 on the pretest based on the responses to the demographics questions (see Appendix B.1). For example, looking at the table, the top-leftmost datum indicates how many students with zero past secondary school and college experience in physical science answered “A” on Question 2. From this raw table, three 2×2 contingency tables for each question were constructed (see Tables 2-7) comparing the accuracy of the responses from students with little to no experience to those from students with some to a lot of experience (N = 103 students). The percent frequencies of each response within each particular bin were also calculated. Tables were previously drawn up comparing the frequencies of individual responses (i.e., A, B, C, D) to each question, but this proved to split the frequencies up too much for the subsequent statistical tests to be informative given the total sample size of my data.

In order to see if transfer occurred on the Post-Exam I survey, I also constructed a raw frequency table for responses to Questions 2 and 4 on the Post-Exam I survey (see Appendix B.2) and used it to draw up 2×2 contingency tables in the same way as for the pretest responses (this time using the responses to the demographics questions on the Post-Exam III survey) (see Tables 8-13). The sample size for the post-test data varied between Questions 2 (N_{Q2} = 67 students) and 4 (N_{Q4} = 53 students) because some students did not answer Question 4. These numbers are smaller than the sample size of the pretest data because, although the total number of responses to the Post-Exam I survey (N_{Q2} = 119 students and N_{Q4} = 100 students) was greater than for the pretest, fewer students responded to the demographics questions in the Post-Exam III survey and so fewer students were counted for the analysis.

Responses to Question 2 on the Pretest According to Total Reported Previous Physical Science Experience							
N (# of students)		103					
	Bin	n (# of students)	$f_{i,Q2}$ (# of responses)	$p_{i,Q2}$ (% of bin)	$f_{c,Q2}$ (# of responses)	$p_{c,Q2}$ (% of bin)	
Total Experience	**0-2**	48	19 / 22.83	39.58% / 47.57%	29 / 25.17	60.42% / 52.43%	
	>2	55	30 / 26.17	54.55% / 47.57%	25 / 28.83	45.45% / 52.43%	
m (# of responses or % of bin)			*49*	*47.57%*	*54*	*52.43%*	
Test for Independence	**Critical Value**	0.05	**Null Hypothesis**	Chosen answer is independent of experience		f_{Q2} is independent of total experience	
	Degrees of Freedom	1	**Alternative Hypothesis**	Chosen answer is dependent on experience		f_{Q2} is dependent on total experience	
	Critical Chi-Squared Value	3.84	**Test Statistic and p Value**	2.30 / 0.129	**Result**	Fail to reject null hypothesis	
Test for Proportionality	**Critical Value**	0.05	**Null Hypothesis**	All groups of students chose the correct answer with the same frequency		$\pi_{c,Q2,0\text{-}2} = \pi_{c,Q2,>2}$	
	Degrees of Freedom	1	**Alternative Hypothesis**	Not all groups of students chose the correct answer with the same frequency		$\pi_{c,Q2,0\text{-}2} \neq \pi_{c,Q2,>2}$	
	Critical Chi-Squared Value	3.841	**Test Statistic and p Value**	1.09 / 0.295	**Result**	Fail to reject null hypothesis	

Table 2. Observed and expected frequencies of correct and incorrect responses to Question 2 on the pretest, organized by total reported secondary school and college experience (in course-years). N is the total sample size, representing the total number of responses to Question 2. n is the sample size of each bin (experience group). f_i is the frequency of incorrect responses and p_i is the percent of students within a bin who responded to Question 2 incorrectly. f_c is the frequency of correct responses to Question 2 (D) and p_c is the percent of students within a bin who responded to Question 2 correctly. The values on the left in a cell are observed values, whereas the values on the right are calculated expected values. m represents the row for the total sample, and is either the total frequency of incorrect/correct responses or the percent of the total sample size that answered incorrectly/correctly, depending on the column. Beneath the data are the critical (alpha) values, degrees of freedom, hypotheses, test statistics, and results of the chi-squared test for independence and chi-squared test for proportionality for Question 2. In order to fit the table on one page, the test statistics and p values have been condensed to a single column; the value on the left is the experimental chi-squared value based on observed data, and the value on the right is the corresponding p value. Total experience was calculated based on answers to Questions 22 and 23 on the pretest. All values are rounded to the hundredth digit, except for the p value which are rounded to the thousandth digit.

Responses to Question 4 on the Pretest According to Total Reported Previous Physical Science Experience							
N (# of students)		103					
	Bin	***n*** (# of students)	$f_{i,Q4}$ (# of responses)	$p_{i,Q4}$ (% of bin)	$f_{c,Q4}$ (# of responses)	$p_{c,Q4}$ (% of bin)	
Total Experience	**0-2**	48	16 / 15.38	33.33% / 32.04%	32 / 32.62	66.67% / 67.96%	
	>2	55	17 / 17.62	30.91% / 32.04%	38 / 37.38	69.09% / 67.96%	
m (# of responses or % of bin)			*33*	*32.04%*	*70*	*67.96%*	
Test for Independence	**Critical Value**	0.05	**Null Hypothesis**	Chosen answer is independent of experience		f_{Q4} is independent of total experience	
	Degrees of Freedom	1	**Alternative Hypothesis**	Chosen answer is dependent on experience		f_{Q4} is dependent on total experience	
	Critical Chi-Squared Value	3.84	**Test Statistic and *p* Value**	0.07 / 0.793	**Result**	Fail to reject null hypothesis	
Test for Proportionality	**Critical Value**	0.05	**Null Hypothesis**	All groups of students chose the correct answer with the same frequency		$\pi_{c,Q4,0\text{-}2} = \pi_{c,Q4,>2}$	
	Degrees of Freedom	1	**Alternative Hypothesis**	Not all groups of students chose the correct answer with the same frequency		$\pi_{c,Q4,0\text{-}2} \neq \pi_{c,Q4,>2}$	
	Critical Chi-Squared Value	3.841	**Test Statistic and *p* Value**	0.02 / 0.882	**Result**	Fail to reject null hypothesis	

Table 3. Observed and expected frequencies of correct and incorrect responses to Question 4 on the pretest, organized by total reported secondary school and college experience (in course-years). N is the total sample size, representing the total number of responses to Question 4. n is the sample size of each bin (experience group). f_i is the frequency of incorrect responses and p_i is the percent of students within a bin who responded to Question 4 incorrectly. f_c is the frequency of correct responses to Question 4 (A) and p_c is the percent of students within a bin who responded to Question 4 correctly. The values on the left in a cell are observed values, whereas the values on the right are calculated expected values. m represents the row for the total sample, and is either the total frequency of incorrect/correct responses or the percent of the total sample size that answered incorrectly/correctly, depending on the column. Beneath the data are the critical (alpha) values, degrees of freedom, hypotheses, test statistics, and results of the chi-squared test for independence and chi-squared test for proportionality for Question 4. In order to fit the table on one page, the test statistics and p values have been condensed to a single column; the value on the left is the experimental chi-squared value based on observed data, and the value on the right is the corresponding p value. Total experience was calculated based on answers to Questions 22 and 23 on the pretest. All values are rounded to the hundredth digit, except for the p value which are rounded to the thousandth digit.

Responses to Question 2 on the Pretest According to Reported Previous Physical Science Experience in Secondary School							
N (# of students)		103					
	Bin	***n*** (# of students)	$f_{i,Q2}$ (# of responses)	$p_{i,Q2}$ (% of bin)	$f_{c,Q2}$ (# of responses)	$p_{c,Q2}$ (% of bin)	
Sec. School Experience	**0-2**	66	30 / 31.40	45.45% / 47.57%	36 / 34.60	54.55% / 52.43%	
	>2	37	19 / 17.60	51.35% / 47.57%	18 / 19.40	48.65% / 52.43%	
m (# of responses or % of bin)			*49*	*47.57%*	*54*	*52.43%*	
Test for Independence	**Critical Value**	0.05	**Null Hypothesis**	Chosen answer is independent of experience		f_{Q2} is independent of sec. school experience	
	Degrees of Freedom	1	**Alternative Hypothesis**	Chosen answer is dependent on experience		f_{Q2} is dependent on sec. school experience	
	Critical Chi-Squared Value	3.84	**Test Statistic and *p* Value**	0.33 / 0.565	**Result**	Fail to reject null hypothesis	
Test for Proportionality	**Critical Value**	0.05	**Null Hypothesis**	All groups of students chose the correct answer with the same frequency		$\pi_{c,Q2,0-2} = \pi_{c,Q2,>2}$	
	Degrees of Freedom	1	**Alternative Hypothesis**	Not all groups of students chose the correct answer with the same frequency		$\pi_{c,Q2,0-2} \neq \pi_{c,Q2,>2}$	
	Critical Chi-Squared Value	3.841	**Test Statistic and *p* Value**	0.16 / 0.692	**Result**	Fail to reject null hypothesis	

Table 4. Observed and expected frequencies of correct and incorrect responses to Question 2 on the pretest, organized by reported secondary school experience (in years). Secondary school experience data comes from the responses to Question 23 on the pretest. See the caption for Table 2 for more information.

Responses to Question 4 on the Pretest According to Reported Previous Physical Science Experience in Secondary School							
N (# of students)		103					
	Bin	***n*** (# of students)	$f_{i,Q4}$ (# of responses)	$p_{i,Q4}$ (% of bin)	$f_{c,Q4}$ (# of responses)	$p_{c,Q4}$ (% of bin)	
Sec. School Experience	**0-2**	66	23 / 21.15	34.85% / 32.04%	43 / 44.85	65.15% / 67.96%	
	>2	37	10 / 11.85	27.03% / 32.04%	27 / 25.15	72.97% / 67.96%	
m (# of responses or % of bin)			*33*	*32.04%*	*70*	*67.96%*	
Test for Independence	**Critical Value**	0.05	**Null Hypothesis**	Chosen answer is independent of experience		f_{Q4} is independent of sec. school experience	
	Degrees of Freedom	1	**Alternative Hypothesis**	Chosen answer is dependent on experience		f_{Q4} is dependent on sec. school experience	
	Critical Chi-Squared Value	3.84	**Test Statistic and *p* Value**	0.67 / 0.414	**Result**	Fail to reject null hypothesis	
Test for Proportionality	**Critical Value**	0.05	**Null Hypothesis**	All groups of students chose the correct answer with the same frequency		$\pi_{c,Q4,0\text{-}2} = \pi_{c,Q4,>2}$	
	Degrees of Freedom	1	**Alternative Hypothesis**	Not all groups of students chose the correct answer with the same frequency		$\pi_{c,Q4,0\text{-}2} \neq \pi_{c,Q4,>2}$	
	Critical Chi-Squared Value	3.841	**Test Statistic and *p* Value**	0.21 / 0.644	**Result**	Fail to reject null hypothesis	

Table 5. Observed and expected frequencies of correct and incorrect responses to Question 4 on the pretest, organized by reported secondary school experience (in years). Secondary school experience data comes from the responses to Question 23 on the pretest. See the caption for Table 3 for more information.

Responses to Question 2 on the Pretest According to Reported Previous Physical Science Experience in College							
N (# of students)		103					
	Bin	***n*** (# of students)	$f_{i,Q2}$ (# of responses)	$p_{i,Q2}$ (% of bin)	$f_{c,Q2}$ (# of responses)	$p_{c,Q2}$ (% of bin)	
College Experience	**0-2**	88	38 / 41.86	43.18% / 47.57%	50 / 46.14	56.82% / 52.43%	
	>2	15	11 / 7.14	73.33% / 47.57%	4 / 7.86	26.67% / 52.43%	
m (# of responses or % of bin)			*49*	*47.57%*	*54*	*52.43%*	
Test for Independence	**Critical Value**	0.05	**Null Hypothesis**	Chosen answer is independent of experience		f_{Q2} is independent of college experience	
	Degrees of Freedom	1	**Alternative Hypothesis**	Chosen answer is dependent on experience		f_{Q2} is dependent on college experience	
	Critical Chi-Squared Value	3.84	**Test Statistic and *p* Value**	4.67 / 0.031	**Result**	Reject null hypothesis	
Test for Proportionality	**Critical Value**	0.05	**Null Hypothesis**	All groups of students chose the correct answer with the same frequency		$\pi_{c,Q2,0\text{-}2} = \pi_{c,Q2,>2}$	
	Degrees of Freedom	1	**Alternative Hypothesis**	Not all groups of students chose the correct answer with the same frequency		$\pi_{c,Q2,0\text{-}2} \neq \pi_{c,Q2,>2}$	
	Critical Chi-Squared Value	3.841	**Test Statistic and *p* Value**	2.22 / 0.136	**Result**	Fail to reject null hypothesis	

Table 6. Observed and expected frequencies of correct and incorrect responses to Question 2 on the pretest, organized by reported college experience (in number of courses). College experience data comes from the responses to Question 22 on the pretest. See the caption for Table 2 for more information.

Responses to Question 4 on the Pretest According to Reported Previous Physical Science Experience in College							
N (# of students)		103					
	Bin	***n*** (# of students)	$f_{i,Q4}$ (# of responses)	$p_{i,Q4}$ (% of bin)	$f_{c,Q4}$ (# of responses)	$p_{c,Q4}$ (% of bin)	
College Experience	**0-2**	88	27 / 28.19	30.68% / 32.04%	61 / 59.81	69.32% / 67.96%	
	>2	15	6 / 4.81	40.00% / 32.04%	9 / 10.19	60.00% / 67.96%	
m (# of responses or % of bin)			*33*	*32.04%*	*70*	*67.96%*	
Test for Independence	**Critical Value**	0.05	**Null Hypothesis**	Chosen answer is independent of experience		f_{Q4} is independent of college experience	
	Degrees of Freedom	1	**Alternative Hypothesis**	Chosen answer is dependent on experience		f_{Q4} is dependent on college experience	
	Critical Chi-Squared Value	3.84	**Test Statistic and *p* Value**	0.51 / 0.475	**Result**	Fail to reject null hypothesis	
Test for Proportionality	**Critical Value**	0.05	**Null Hypothesis**	All groups of students chose the correct answer with the same frequency		$\pi_{c,Q4,0\text{-}2} = \pi_{c,Q4,>2}$	
	Degrees of Freedom	1	**Alternative Hypothesis**	Not all groups of students chose the correct answer with the same frequency		$\pi_{c,Q4,0\text{-}2} \neq \pi_{c,Q4,>2}$	
	Critical Chi-Squared Value	3.841	**Test Statistic and *p* Value**	0.16 / 0.685	**Result**	Fail to reject null hypothesis	

Table 7. *O* bserved and expected frequencies of correct and incorrect responses to Question 2 on the pretest, organized by reported college experience (in number of courses). College experience data comes from the responses to Question 22 on the pretest. See the caption for Table 3 for more information.

Responses to Question 2 on the Post-Exam I Survey According to Total Reported Previous Physical Science Experience						
N (# of students)		56				
	Bin	n (# of students)	$f_{i,Q2}$ (# of responses)	$p_{i,Q2}$ (% of bin)	$f_{c,Q2}$ (# of responses)	$p_{c,Q2}$ (% of bin)
Total Experience	**0-2**	27	14 / 9.16	51.85% / 33.93%	13 / 17.84	48.15% / 66.07%
	>2	29	5 / 9.84	17.24% / 33.93%	24 / 19.16	82.76% / 66.07%
m (# of responses or % of bin)			*19*	*33.93%*	*37*	*66.07%*
Test for Independence	**Critical Value**	0.05	**Null Hypothesis**	Chosen answer is independent of experience		f_{Q2} is independent of total experience
	Degrees of Freedom	1	**Alternative Hypothesis**	Chosen answer is dependent on experience		f_{Q2} is dependent on total experience
	Critical Chi-Squared Value	3.84	**Test Statistic and p Value**	7.47 / 0.006	**Result**	Reject null hypothesis
Test for Proportionality	**Critical Value**	0.05	**Null Hypothesis**	All groups of students chose the correct answer with the same frequency		$\pi_{c,Q2,0\text{-}2} = \pi_{c,Q2,>2}$
	Degrees of Freedom	1	**Alternative Hypothesis**	Not all groups of students chose the correct answer with the same frequency		$\pi_{c,Q2,0\text{-}2} \neq \pi_{c,Q2,>2}$
	Critical Chi-Squared Value	3.841	**Test Statistic and p Value**	2.53 / 0.111	**Result**	Fail to reject null hypothesis

Table 8. Observed and expected frequencies of correct and incorrect responses to Question 2 on the Post-Exam I survey, organized by total reported secondary school and college experience (in course-years). Total experience was calculated based on answers to Questions 17 and 18 on the Post-Exam III survey. See the caption for Table 2 for more information.

Responses to Question 4 on the Post-Exam I Survey According to Total Reported Previous Physical Science Experience							
N (# of students)		53					
	Bin	***n*** (# of students)	$f_{i,Q4}$ (# of responses)	$p_{i,Q4}$ (% of bin)	$f_{c,Q4}$ (# of responses)	$p_{c,Q4}$ (% of bin)	
Total Experience	**0-2**	26	14 / 13.74	53.85% / 52.83%	12 / 12.26	46.15% / 47.17%	
	>2	27	14 / 14.26	51.85% / 52.83%	13 / 12.74	48.15% / 47.17%	
m (# of responses or % of bin)			*28*	*52.83%*	*25*	*47.17%*	
Test for Independence	**Critical Value**	0.05	**Null Hypothesis**	Chosen answer is independent of experience		f_{Q4} is independent of total experience	
	Degrees of Freedom	1	**Alternative Hypothesis**	Chosen answer is dependent on experience		f_{Q4} is dependent on total experience	
	Critical Chi-Squared Value	3.84	**Test Statistic and *p* Value**	0.02 / 0.885	**Result**	Fail to reject null hypothesis	
Test for Proportionality	**Critical Value**	0.05	**Null Hypothesis**	All groups of students chose the correct answer with the same frequency		$\pi_{c,Q4,0\text{-}2} = \pi_{c,Q4,>2}$	
	Degrees of Freedom	1	**Alternative Hypothesis**	Not all groups of students chose the correct answer with the same frequency		$\pi_{c,Q4,0\text{-}2} \neq \pi_{c,Q4,>2}$	
	Critical Chi-Squared Value	3.841	**Test Statistic and *p* Value**	0.01 / 0.916	**Result**	Fail to reject null hypothesis	

Table 9. Observed and expected frequencies of correct and incorrect responses to Question 4 on the Post-Exam I survey, organized by total reported secondary school and college experience (in course-years). Total experience was calculated based on answers to Questions 17 and 18 on the Post-Exam III survey. See the caption for Table 3 for more information.

Responses to Question 2 on the Post-Exam I Survey According to Reported Previous Physical Science Experience in Secondary School							
N (# of students)		56					
	Bin	***n*** (# of students)	$f_{i,Q2}$ (# of responses)	$p_{i,Q2}$ (% of bin)	$f_{c,Q2}$ (# of responses)	$p_{c,Q2}$ (% of bin)
Sec. School Experience	**0-2**	42	18 / 14.25	42.86% / 33.93%	24 / 27.75	57.14% / 66.07%
	>2	14	1 / 4.75	7.14% / 33.93%	13 / 9.25	92.86% / 66.07%
m (# of responses or % of bin)			*19*	*33.93%*	*37*	*66.07%*
Test for Independence	**Critical Value**	0.05	**Null Hypothesis**	Chosen answer is independent of experience		f_{Q2} is independent of sec. school experience
	Degrees of Freedom	1	**Alternative Hypothesis**	Chosen answer is dependent on experience		f_{Q2} is dependent on sec. school experience
	Critical Chi-Squared Value	3.84	**Test Statistic and *p* Value**	5.97 / 0.015	**Result**	Reject null hypothesis
Test for Proportionality	**Critical Value**	0.05	**Null Hypothesis**	All groups of students chose the correct answer with the same frequency		$\pi_{c,Q2,0\text{-}2} = \pi_{c,Q2,>2}$
	Degrees of Freedom	1	**Alternative Hypothesis**	Not all groups of students chose the correct answer with the same frequency		$\pi_{c,Q2,0\text{-}2} \neq \pi_{c,Q2,>2}$
	Critical Chi-Squared Value	3.841	**Test Statistic and *p* Value**	2.03 / 0.154	**Result**	Fail to reject null hypothesis

Table 10. Observed and expected frequencies of correct and incorrect responses to Question 2 on the Post-Exam I survey, organized by reported secondary school experience (in years). Secondary school experience data comes from the responses to Question 18 on the Post-Exam III survey. See the caption for Table 2 for more information.

Responses to Question 4 on the Post-Exam I Survey According to Reported Previous Physical Science Experience in Secondary School						
N (# of students)		53				
	Bin	n (# of students)	$f_{i,Q4}$ (# of responses)	$p_{i,Q4}$ (% of bin)	$f_{c,Q4}$ (# of responses)	$p_{c,Q4}$ (% of bin)
Sec. School Experience	**0-2**	39	20 / 20.6	51.28% / 52.83%	19 / 18.4	48.72% / 47.17%
	>2	14	8 / 7.4	57.14% / 52.83%	6 / 6.6	42.86% / 47.17%
m (# of responses or % of bin)			*28*	*52.83%*	*25*	*47.17%*
Test for Independence	**Critical Value**	0.05	**Null Hypothesis**	Chosen answer is independent of experience		f_{Q4} is independent of sec. school experience
	Degrees of Freedom	1	**Alternative Hypothesis**	Chosen answer is dependent on experience		f_{Q4} is dependent on sec. school experience
	Critical Chi-Squared Value	3.84	**Test Statistic and p Value**	0.14 / 0.706	**Result**	Fail to reject null hypothesis
Test for Proportionality	**Critical Value**	0.05	**Null Hypothesis**	All groups of students chose the correct answer with the same frequency		$\pi_{c,Q4,0\text{-}2} = \pi_{c,Q4,>2}$
	Degrees of Freedom	1	**Alternative Hypothesis**	Not all groups of students chose the correct answer with the same frequency		$\pi_{c,Q4,0\text{-}2} \neq \pi_{c,Q4,>2}$
	Critical Chi-Squared Value	3.841	**Test Statistic and p Value**	0.08 / 0.784	**Result**	Fail to reject null hypothesis

Table 11. Observed and expected frequencies of correct and incorrect responses to Question 4 on the Post-Exam I survey, organized by reported secondary school experience (in years). Secondary school experience data comes from the responses to Question 18 on the Post-Exam III survey. See the caption for Table 3 for more information.

Responses to Question 2 on the Post-Exam I Survey According to Reported Previous Physical Science Experience in College							
N (# of students)		56					
	Bin	n (# of students)	$f_{i,Q2}$ (# of responses)	$p_{i,Q2}$ (% of bin)	$f_{c,Q2}$ (# of responses)	$p_{c,Q2}$ (% of bin)
College Experience	0-2	47	18 / 15.95	38.30% / 33.93%	29 / 31.05	61.70% / 66.07%
	>2	9	1 / 3.05	11.11% / 33.93%	8 / 5.95	88.89% / 66.07%
m (# of responses or % of bin)			*19*	*33.93%*	*37*	*66.07%*
Test for Independence	**Critical Value**	0.05	**Null Hypothesis**	Chosen answer is independent of experience		f_{Q2} is independent of college experience
	Degrees of Freedom	1	**Alternative Hypothesis**	Chosen answer is dependent on experience		f_{Q2} is dependent on college experience
	Critical Chi-Squared Value	3.84	**Test Statistic and p Value**	2.49 / 0.115	**Result**	Fail to reject null hypothesis
Test for Proportionality	**Critical Value**	0.05	**Null Hypothesis**	All groups of students chose the correct answer with the same frequency		$\pi_{c,Q2,0\text{-}2} = \pi_{c,Q2,>2}$
	Degrees of Freedom	1	**Alternative Hypothesis**	Not all groups of students chose the correct answer with the same frequency		$\pi_{c,Q2,0\text{-}2} \neq \pi_{c,Q2,>2}$
	Critical Chi-Squared Value	3.841	**Test Statistic and p Value**	0.84 / 0.358	**Result**	Fail to reject null hypothesis

Table 12. Observed and expected frequencies of correct and incorrect responses to Question 2 on the Post-Exam I survey, organized by reported college experience (in number of courses). College experience data comes from the responses to Question 17 on the Post-Exam III survey. See the caption for Table 2 for more information.

Responses to Question 4 on the Post-Exam I Survey According to Reported Previous Physical Science Experience in College							
N (# of students)		53					
	Bin	***n*** (# of students)	$f_{i,Q4}$ (# of responses)	$p_{i,Q4}$ (% of bin)	$f_{c,Q4}$ (# of responses)	$p_{c,Q4}$ (% of bin)	
College Experience	**0-2**	44	24 / 23.25	54.55% / 52.83%	20 / 20.75	45.45% / 47.17%	
	>2	9	4 / 4.75	44.44% / 52.83%	5 / 4.25	55.56% / 47.17%	
m (# of responses or % of bin)			*28*	*52.83%*	*25*	*47.17%*	
Test for Independence	**Critical Value**	0.05	**Null Hypothesis**	Chosen answer is independent of experience		f_{Q4} is independent of college experience	
	Degrees of Freedom	1	**Alternative Hypothesis**	Chosen answer is dependent on experience		f_{Q4} is dependent on college experience	
	Critical Chi-Squared Value	3.84	**Test Statistic and *p* Value**	0.31 / 0.580	**Result**	Fail to reject null hypothesis	
Test for Proportionality	**Critical Value**	0.05	**Null Hypothesis**	All groups of students chose the correct answer with the same frequency		$\pi_{c,Q4,0\text{-}2} = \pi_{c,Q4,>2}$	
	Degrees of Freedom	1	**Alternative Hypothesis**	Not all groups of students chose the correct answer with the same frequency		$\pi_{c,Q4,0\text{-}2} \neq \pi_{c,Q4,>2}$	
	Critical Chi-Squared Value	3.841	**Test Statistic and *p* Value**	0.16 / 0.688	**Result**	Fail to reject null hypothesis	

Table 13. Observed and expected frequencies of correct and incorrect responses to Question 4 on the Post-Exam I survey, organized by reported college experience (in number of courses). College experience data comes from the responses to Question 17 on the Post-Exam III survey. See the caption for Table 3 for more information.

It is once more important to stress that Appendix B.2 and Tables 8-13 are for just the post-tests *only*. The tables reflect the performance of a group of students who partially overlapped with the students who took the pretest but which also includes some students who did not take the pretest and excludes some students who took the pretest but dropped the course before taking the Post-Exam III survey.

In order to get a sense of how just the students who took the pretest fared during the course, a set of three additional raw frequency tables were constructed that compared how frequently students with a particular amount of (total, secondary school, or college) experience who chose a specific responses on Questions 2 and 4 of the pretest selected particular answers to Questions 2 and 4 on the Post-Exam I survey (see Appendix C.1-C.3). For example, in Appendix C.1, the top-leftmost datum indicates the number of students with no physical science experience at the beginning of the semester who chose answer "A" on the pretest and answer "A" on the Post-Exam I survey for Question 2. This may alternatively be interpreted as the number of students with no physical science experience at the beginning of the semester who originally chose answer "A" on Question 2 of the pretest and did not subsequently change their answer on the Post-Exam I survey. The utility of these tables was that it accounted for the inconsistencies between reported experience on the pretest and the Post-Exam III survey by only basing physical science experience on responses to the demographics questions on the pretest. Because students who did not make it through the course long enough to take the Post-Exam III survey were still able to be counted towards the analysis, these tables showed an improved sample size compared to the post-test analysis: 74 students for Question 2 and 64 students for Question 4. Six 4×3 contingency tables were drawn up, two for each type of experience (total, just secondary school, or just college), but experience this time was compared to whether a student changed their answer on the Post-Exam I survey (either to an incorrect answer or the correct answer) or kept their answer the same (whether it was correct or incorrect) (see Tables 14-19).

For the multiple-choice data, two chi-squared tests were performed for each table in order to test whether or not students were engaging in far transfer: a test for independence and a test for proportionality (Pearson, 1900) (see Tables 2-19). For Tables 2-13, this study posits that, if students are engaging in far transfer, the responses from students with physical science background should be different from those from students with little to no physical science background. The chi-squared tests for independence for the pretest analyses and post-test analyses were performed to see whether previous background knowledge about physical science had an effect on which responses students selected for Questions 2 and 4. The null hypothesis is that the answers students chose for a particular question were independent of their physical science experience (mathematically, Pr{Incorrect|0-2} = Pr{Correct|0-2}). The alternative hypothesis is that the answers students chose for a particular question were dependent on their amount of physical science experience (Pr{Incorrect|0-2} ≠ Pr{Correct|0-2}). A critical value of $\alpha = 0.05$ was chosen because that is typical for educational research, and there was 1 degree of freedom for all tables ([2 – 1][2 – 1]). The chi-squared tests for proportionality for the pretest analyses and post-test analyses were performed in order to see whether there was a statistically significant difference between the number of correct answers depending on physical science experience. The null hypothesis for Question 2 is that all students chose the correct answer “D” with the same frequency regardless of their physical science experience ($\pi_{correct,0-2} = \pi_{correct,>2}$). The null hypothesis for Question 4 is that all students chose the correct answer “A” with the same frequency regardless of their physical science experience. The alternative hypotheses for each question are that physical science experience made a difference between how many students from each bin selected the correct answer ($\pi_{correct,0-2} \neq \pi_{correct,>2}$). Like the other chi-squared tests, a critical value of α = 0.05 was once again chosen. There was 1 degree of freedom for all six tables (2 – 1).

For Tables 14-19, this study posits that, if transfer did not necessarily initially occur but influenced how students learned new material, there should be a difference between how students with varying past physical science experience changed their answers to Questions 2 or 4 on the Post-Exam I survey compared to what they responded on the pretest. The critical values for the tests were identical to those used for the pretest analyses and post-test analyses. The chi-squared tests for independence had 3 degrees of freedom ([4 – 1][2 – 1]) and the chi-squared tests for proportionality had 1 degree of freedom ([2 – 1][2 – 1]). The null hypotheses for the tests for independence were that whether or not a student answered a particular question differently on the Post-Exam I survey was statistically independent of their past physical science experience before taking the course (mathematically, Pr{Remains Incorrect|0-2} = Pr{Changes to Incorrect|0-2} = Pr{Changes to Correct|0-2} = Pr{Remains Correct|0-2}). The alternative hypotheses were that the two variables were statistically dependent (mathematically, at least one of the aforementioned probabilities was not equal to the rest). The null hypotheses for the tests for proportionality were that past physical science experience did not make a statistically significant difference in how many students changed an incorrect answer on the pretest to the correct answer on the Post-Exam I survey ($\pi_{changes\ to\ correct,0-2} = \pi_{changes\ to\ correct,>2}$). The alternative hypotheses were that not all experience groups changed their previously incorrect answers to correct ones on the Post-Exam I survey ($\pi_{changes\ to\ correct,0-2} \neq \pi_{changes\ to\ correct,>2}$).

Change in Responses to Question 2 Between the Pretest and the Post-Exam I Survey, According to Total Reported Physical Science Experience										
N (# of students)		74								
	Bin	*n* (# of students)	$f_{remains\ incorrect,Q2}$ (# of responses)	$p_{remains\ incorrect,Q2}$ (% of bin)	$f_{changes\ to\ incorrect,Q2}$ (# of responses)	$p_{changes\ to\ incorrect,Q2}$ (% of bin)	$f_{changes\ to\ correct,Q2}$ (# of responses)	$p_{changes\ to\ correct,Q2}$ (% of bin)	$f_{remains\ correct,Q2}$ (# of responses)	$p_{remains\ correct,Q2}$ (% of bin)
Total Experience	**0-2**	35	8 / 8.51	22.86% / 24.32%	6 / 4.73	17.14% / 13.51%	5 / 8.04	14.29% / 22.97%	16 / 13.72	45.71% / 39.19%
	>2	39	10 / 9.49	25.64% / 24.32%	4 / 5.27	10.26% / 13.51%	12 / 8.96	30.77% / 22.97%	13 / 15.28	33.33% / 39.19%
m (# of responses or % of bin)			*18*	*24.32%*	*10*	*13.51%*	*17*	*22.97%*	*29*	*39.19%*
Test for Independence	**Critical Value**	0.05	**Null Hypothesis**	Whether or not students changed their answer is independent of experience					f_{Q2} is independent of total experience	
	Degrees of Freedom	3	**Alternative Hypothesis**	Whether or not students changed their answer is dependent on experience					f_{Q2} is dependent on total experience	
	Critical Chi-Squared Value	7.81	**Test Statistic**	3.61	***p* Value**	0.307	**Result**	Fail to reject null hypothesis		
Test for Proportionality	**Critical Value**	0.05	**Null Hypothesis**	All groups of students who originally had an incorrect answer on the pre-test changed to Answer D on the post-test with the same frequency					$\pi_{changes\ to\ correct,Q2,0\text{-}2} = \pi_{changes\ to\ correct,Q2,>2}$	
	Degrees of Freedom	1	**Alternative Hypothesis**	Not all groups of students who originally had an incorrect answer on the pre-test changed to Answer D on the post-test with the same frequency					$\pi_{changes\ to\ correct,Q2,0\text{-}2} \neq \pi_{changes\ to\ correct,Q2,>2}$	
	Critical Chi-Squared Value	3.84	**Test Statistic**	2.18	***p* Value**	0.140	**Result**	Fail to reject null hypothesis		

Table 14. Observed and expected frequencies at which students did or did not change their response to Question 2 between the pretest and the Post-Exam I survey, organized by total reported secondary school and college experience (in course-years). $f_{remains\ incorrect}$ is the frequency that students who answered Question 2 incorrectly on the pretest also answered Question 2 incorrectly on the Post-Exam I survey. $p_{remains\ incorrect}$ is the percent of students within a bin who answered Question 2 incorrectly on the pretest also answered Question 2 incorrectly on the Post-Exam I survey. The "changes to incorrect" columns represent students who originally answered Question 2 on the pretest correctly but then changed their answer to an incorrect one on the Post-Exam I survey. The "changes to correct" columns represent students who originally answered Question 2 on the pretest incorrectly but then changed their answer to a correct one on the Post-Exam I survey. The "remains correct" columns represent students who originally answered Question 2 on the pretest correctly and did not change their answer on the Post-Exam I survey. Total experience was calculated based on answers to Questions 22 and 23 on the pretest and Questions 17 and 18 on the Post-Exam III survey. See the caption for Table 2 for more information.

Change in Responses to Question 4 Between the Pretest and the Post-Exam I Survey, According to Total Reported Physical Science Experience											
N (# of students)		64									
	Bin	***n*** (# of students)	$f_{remains\ incorrect,Q4}$ (# of responses)	$p_{remains\ incorrect,Q4}$ (% of bin)	$f_{changes\ to\ incorrect,Q4}$ (# of responses)	$p_{changes\ to\ incorrect,Q4}$ (% of bin)	$f_{changes\ to\ correct,Q4}$ (# of responses)	$p_{changes\ to\ correct,Q4}$ (% of bin)	$f_{remains\ correct,Q4}$ (# of responses)	$p_{remains\ correct,Q4}$ (% of bin)	
Total Experience	**0-2**	27	3 / 4.64	11.11% / 17.19%	8 / 7.17	29.63% / 26.56%	4 / 3.8	14.81% / 14.06%	12 / 11.39	44.44% / 42.19%	
	>2	37	8 / 6.36	21.62% / 17.19%	9 / 9.83	24.32% / 26.56%	5 / 5.2	13.51% / 14.06%	15 / 15.61	40.54% / 42.19%	
m (# of responses or % of bin)			*11*	*17.19%*	*17*	*26.56%*	*9*	*14.06%*	*27*	*42.19%*	
Test for Independence	**Critical Value**	0.05	**Null Hypothesis**	Whether or not students changed their answer is independent of experience					f_{Q4} is independent of total experience		
	Degrees of Freedom	3	**Alternative Hypothesis**	Whether or not students changed their answer is dependent on experience					f_{Q4} is dependent on total experience		
	Critical Chi-Squared Value	7.81	**Test Statistic**	1.24	***p* Value**	0.743	**Result**	Fail to reject null hypothesis			
Test for Proportionality	**Critical Value**	0.05	**Null Hypothesis**	All groups of students who originally had an incorrect answer on the pre-test changed to Answer A on the post-test with the same frequency					$\pi_{changes\ to\ correct,Q4,0\text{-}2} = \pi_{changes\ to\ correct,Q4,>2}$		
	Degrees of Freedom	1	**Alternative Hypothesis**	Not all groups of students who originally had an incorrect answer on the pre-test changed to Answer A on the post-test with the same frequency					$\pi_{changes\ to\ correct,Q4,0\text{-}2} \neq \pi_{changes\ to\ correct,Q4,>2}$		
	Critical Chi-Squared Value	3.84	**Test Statistic**	0.02	***p* Value**	0.891	**Result**	Fail to reject null hypothesis			

Table 15. Observed and expected frequencies at which students did or did not change their response to Question 4 between the pretest and the Post-Exam I survey, organized by total reported secondary school and college experience (in course-years). $f_{remains\ incorrect}$ is the frequency that students who answered Question 4 incorrectly on the pretest also answered Question 4 incorrectly on the Post-Exam I survey. $p_{remains\ incorrect}$ is the percent of students within a bin who answered Question 4 incorrectly on the pretest also answered Question 4 incorrectly on the Post-Exam I survey. The "changes to incorrect" columns represent students who originally answered Question 4 on the pretest correctly but then changed their answer to an incorrect one on the Post-Exam I survey. The "changes to correct" columns represent students who originally answered Question 4 on the pretest incorrectly but then changed their answer to a correct one on the Post-Exam I survey. The "remains correct" columns represent students who originally answered Question 4 on the pretest correctly and did not change their answer on the Post-Exam I survey. Total experience was calculated based on answers to Questions 22 and 23 on the pretest and Questions 17 and 18 on the Post-Exam III survey. See the caption for Table 3 for more information.

Change in Responses to Question 2 Between the Pretest and the Post-Exam I Survey, According to Reported Physical Science Experience in Secondary School										
N (# of students)		74								
	Bin	n (# of students)	$f_{remains\ incorrect,Q2}$ (# of responses)	$p_{remains\ incorrect,Q2}$ (% of bin)	$f_{changes\ to\ incorrect,Q2}$ (# of responses)	$p_{changes\ to\ incorrect,Q2}$ (% of bin)	$f_{changes\ to\ correct,Q2}$ (# of responses)	$p_{changes\ to\ correct,Q2}$ (% of bin)	$f_{remains\ correct,Q2}$ (# of responses)	$p_{remains\ correct,Q2}$ (% of bin)
Sec. School Experience	**0-2**	46	10 / 11.19	21.74% / 24.32%	6 / 6.22	13.04% / 13.51%	11 / 10.57	23.91% / 22.97%	19 / 18.03	41.30% / 39.19%
	>2	28	8 / 6.81	28.57% / 24.32%	4 / 3.78	14.29% / 13.51%	6 / 6.43	21.43% / 22.97%	10 / 10.97	35.71% / 39.19%
m (# of responses or % of bin)			*18*	*24.32%*	*10*	*13.51%*	*17*	*22.97%*	*29*	*39.19%*
Test for Independence	**Critical Value**	0.05	**Null Hypothesis**	Whether or not students changed their answer is independent of experience					f_{Q2} is independent of secondary school experience	
	Degrees of Freedom	3	**Alternative Hypothesis**	Whether or not students changed their answer is dependent on experience					f_{Q2} is dependent on secondary school experience	
	Critical Chi-Squared Value	7.81	**Test Statistic**	0.54	***p* Value**	0.910	**Result**	Fail to reject null hypothesis		
Test for Proportionality	**Critical Value**	0.05	**Null Hypothesis**	All groups of students who originally had an incorrect answer on the pre-test changed to Answer D on the post-test with the same frequency					$\pi_{changes\ to\ correct,Q2,0\text{-}2} = \pi_{changes\ to\ correct,Q2,>2}$	
	Degrees of Freedom	1	**Alternative Hypothesis**	Not all groups of students who originally had an incorrect answer on the pre-test changed to Answer D on the post-test with the same frequency					$\pi_{changes\ to\ correct,Q2,0\text{-}2} \neq \pi_{changes\ to\ correct,Q2,>2}$	
	Critical Chi-Squared Value	3.84	**Test Statistic**	0.05	***p* Value**	0.829	**Result**	Fail to reject null hypothesis		

Table 16. Observed and expected frequencies at which students did or did not change their response to Question 2 between the pretest and the Post-Exam I survey, organized by reported secondary school experience (in years). Secondary school experience is based on the responses to Question 23 on the pretest and Question 18 on the Post-Exam III survey. See the captions for Tables 2 and 14 for more information.

Change in Responses to Question 4 Between the Pretest and the Post-Exam I Survey, According to Reported Physical Science Experience in Secondary School										
N (# of students)		64								
	Bin	n (# of students)	$f_{i,Q4}$ (# of responses)	$p_{i,Q4}$ (% of bin)	$f_{changes\ to\ incorrect,Q4}$ (# of responses)	$p_{changes\ to\ incorrect,Q4}$ (% of bin)	$f_{changes\ to\ correct,Q4}$ (# of responses)	$p_{changes\ to\ correct,Q4}$ (% of bin)	$f_{c,Q4}$ (# of responses)	$p_{c,Q4}$ (% of bin)
Sec. School Experience	0-2	38	5 / 6.53	13.16% / 17.19%	10 / 10.09	26.32% / 26.56%	7 / 5.34	18.42% / 14.06%	16 / 16.03	42.11% / 42.19%
	>2	26	6 / 4.47	23.08% / 17.19%	7 / 6.91	26.92% / 26.56%	2 / 3.66	7.69% / 14.06%	11 / 10.97	42.31% / 42.19%
m (# of responses or % of bin)			*11*	*17.19%*	*17*	*26.56%*	*9*	*14.06%*	*27*	*42.19%*
Test for Independence	Critical Value	0.05	Null Hypothesis	Whether or not students changed their answer is independent of experience					f_{Q4} is independent of secondary school experience	
	Degrees of Freedom	3	Alternative Hypothesis	Whether or not students changed their answer is dependent on experience					f_{Q4} is dependent on secondary school experience	
	Critical Chi-Squared Value	7.81	Test Statistic	2.15	p Value	0.542	Result	Fail to reject null hypothesis		
Test for Proportionality	Critical Value	0.05	Null Hypothesis	All groups of students who originally had an incorrect answer on the pre-test changed to Answer A on the post-test with the same frequency					$\pi_{changes\ to\ correct,Q4,0\text{-}2} = \pi_{changes\ to\ correct,Q4,>2}$	
	Degrees of Freedom	1	Alternative Hypothesis	Not all groups of students who originally had an incorrect answer on the pre-test changed to Answer A on the post-test with the same frequency					$\pi_{changes\ to\ correct,Q4,0\text{-}2} \neq \pi_{changes\ to\ correct,Q4,>2}$	
	Critical Chi-Squared Value	3.84	Test Statistic	1.26	p Value	0.261	Result	Fail to reject null hypothesis		

Table 17. Observed and expected frequencies at which students did or did not change their response to Question 4 between the pretest and the Post-Exam I survey, organized by reported secondary school experience (in years). Secondary school experience is based on the responses to Question 23 on the pretest and Question 18 on the Post-Exam III survey. See the captions for Tables 3 and 15 for more information.

Change in Responses to Question 2 Between the Pretest and the Post-Exam I Survey, According to Reported Physical Science Experience in College										
N (# of students)		74								
	Bin	***n*** (# of students)	$f_{remains\ incorrect,Q2}$ (# of responses)	$p_{remains\ incorrect,Q2}$ (% of bin)	$f_{changes\ to\ incorrect,Q2}$ (# of responses)	$p_{changes\ to\ incorrect,Q2}$ (% of bin)	$f_{changes\ to\ correct,Q2}$ (# of responses)	$p_{changes\ to\ correct,Q2}$ (% of bin)	$f_{remains\ correct,Q2}$ (# of responses)	$p_{remains\ correct,Q2}$ (% of bin)
College Experience	**0-2**	65	16 / 15.81	24.62% / 24.32%	10 / 8.78	15.38% / 13.51%	12 / 14.93	18.46% / 22.97%	27 / 25.47	41.54% / 39.19%
	>2	9	2 / 2.19	22.22% / 24.32%	0 / 1.22	0.00% / 13.51%	5 / 2.07	55.56% / 22.97%	2 / 3.53	22.22% / 39.19%
m (# of responses or % of bin)			*18*	*24.32%*	*10*	*13.51%*	*17*	*22.97%*	*29*	*39.19%*
Test for Independence	**Critical Value**	0.05	**Null Hypothesis**	Whether or not students changed their answer is independent of experience					f_{Q2} is independent of college experience	
	Degrees of Freedom	3	**Alternative Hypothesis**	Whether or not students changed their answer is dependent on experience					f_{Q2} is dependent on college experience	
	Critical Chi-Squared Value	7.81	**Test Statistic**	6.89	***p* Value**	0.075	**Result**	Fail to reject null hypothesis		
Test for Proportionality	**Critical Value**	0.05	**Null Hypothesis**	All groups of students who originally had an incorrect answer on the pre-test changed to Answer D on the post-test with the same frequency					$\pi_{changes\ to\ correct,Q2,0\text{-}2} = \pi_{changes\ to\ correct,Q2,>2}$	
	Degrees of Freedom	1	**Alternative Hypothesis**	Not all groups of students who originally had an incorrect answer on the pre-test changed to Answer D on the post-test with the same frequency					$\pi_{changes\ to\ correct,Q2,0\text{-}2} \neq \pi_{changes\ to\ correct,Q2,>2}$	
	Critical Chi-Squared Value	3.84	**Test Statistic**	4.73	***p* Value**	0.030	**Result**	Reject null hypothesis		

Table 18. Observed and expected frequencies that compares student responses to Question 2 on the pretest to their responses to Question 2 on the Post-Exam I survey, organized by reported college experience (in courses). College experience is based on the responses to Question 22 on the pretest and Question 17 on the Post-Exam III survey. See the captions for Tables 2 and 14 for more information.

Change in Responses to Question 4 Between the Pretest and the Post-Exam I Survey, According to Reported Physical Science Experience in College										
N (# of students)		64								
	Bin	***n*** (# of students)	$f_{i,Q4}$ (# of responses)	$p_{i,Q4}$ (% of bin)	$f_{changes\ to\ incorrect,Q4}$ (# of responses)	$p_{changes\ to\ incorrect,Q4}$ (% of bin)	$f_{changes\ to\ correct,Q4}$ (# of responses)	$p_{changes\ to\ correct,Q4}$ (% of bin)	$f_{c,Q4}$ (# of responses)	$p_{c,Q4}$ (% of bin)
College Experience	**0-2**	55	9 / 9.45	16.36% / 17.19%	16 / 14.61	29.09% / 26.56%	7 / 7.73	12.73% / 14.06%	23 / 23.2	41.82% / 42.19%
	>2	9	2 / 1.55	22.22% / 17.19%	1 / 2.39	11.11% / 26.56%	2 / 1.27	22.22% / 14.06%	4 / 3.8	44.44% / 42.19%
m (# of responses or % of bin)			*11*	*17.19%*	*17*	*26.56%*	*9*	*14.06%*	*27*	*42.19%*
Test for Independence	**Critical Value**	0.05	**Null Hypothesis**	Whether or not students changed their answer is independent of experience					f_{Q4} is independent of college experience	
	Degrees of Freedom	3	**Alternative Hypothesis**	Whether or not students changed their answer is dependent on experience					f_{Q4} is dependent on college experience	
	Critical Chi-Squared Value	7.81	**Test Statistic**	1.60	***p* Value**	0.658	**Result**	Fail to reject null hypothesis		
Test for Proportionality	**Critical Value**	0.05	**Null Hypothesis**	All groups of students who originally had an incorrect answer on the pre-test changed to Answer A on the post-test with the same frequency					$\pi_{changes\ to\ correct,Q4,0-2} = \pi_{changes\ to\ correct,Q4,>2}$	
	Degrees of Freedom	1	**Alternative Hypothesis**	Not all groups of students who originally had an incorrect answer on the pre-test changed to Answer A on the post-test with the same frequency					$\pi_{changes\ to\ correct,Q4,0-2} \neq \pi_{changes\ to\ correct,Q4,>2}$	
	Critical Chi-Squared Value	3.84	**Test Statistic**	0.50	***p* Value**	0.482	**Result**	Fail to reject null hypothesis		

Table 19. Observed and expected frequencies that compares student responses to Question 4 on the pretest to their responses to Question 2 on the Post-Exam I survey, organized by reported college experience (in courses). College experience is based on the responses to Question 22 on the pretest and Question 17 on the Post-Exam III survey. See the captions for Tables 3 and 15 for more information.

Results of Multiple-Choice Analysis

I could not rule out the null hypotheses of any pretest analysis except for the chi-squared test for independence for Question 2 that I performed on Table 6 (see Tables 2-7 for chi-squared values and *p* values). This means that it cannot be ruled out that the answers that students selected on the pretest were independent of past physical science experience in secondary school nor that the answers were independent of total past physical science experience. Student answers to Question 4 on Table 7 could not be ruled out as being independent of college experience, but I can reject the null hypothesis and conclude that statistically, student answers to Question 2 were dependent on past physical science experience in college. However, as previously mentioned, I still could not rule out the possibility that any experience group chose the correct answer to Question 2 more or less frequently than would be expected if past college experience had no effect.

The results of my chi-squared test analysis of the effect of college experience on the pretest responses to Questions 2 and 4 were also true for the effects of secondary school experience and total experience on the Post-Exam I survey responses (see Tables 8-11 for chi-squared values and *p* values). It can be ruled out that whether a student was correct or incorrect on Question 2 of the Post-Exam I survey was independent of their secondary school experience as well as independent of their total previous experience with formal physical science education. However, I could not rule out the null hypotheses for those analyses' respective chi-squared tests for proportionality. No null hypothesis could be ruled out when looking at just college experience (see Tables 12 and 13 for chi-squared values and *p* values).

Graphs comparing student total physical science experience, secondary school physical science experience, and college physical science experience to the percent frequency of correct answers were produced for both the pretest data and the post-test data (see Figures 2-5).

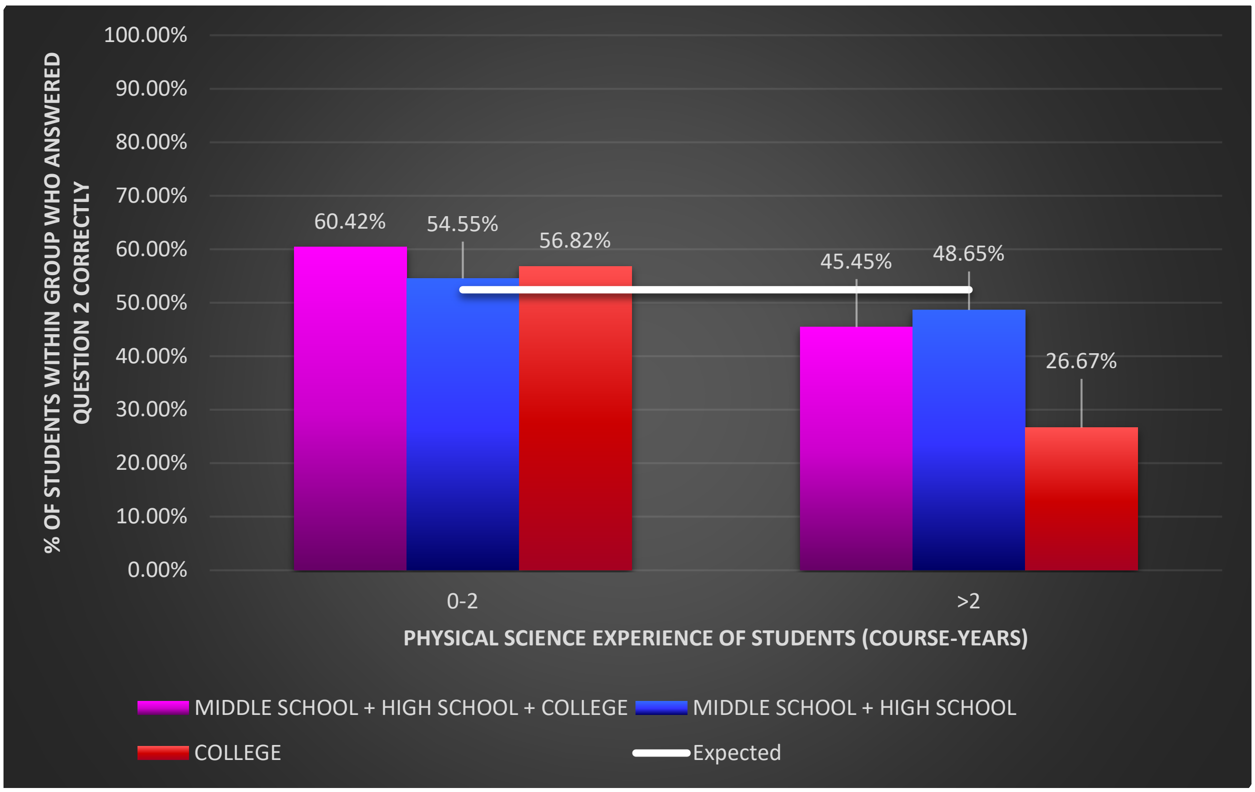

Figure 2. Percentage of students in each experience group who correctly answered “D” on Question 2 of the pretest. The purple bar is based on total previous physical science experience; the blue bar is based on just previous physical science experience in secondary school; the red bar is based on just previous physical science experience in college. The white line represents the expected percentage. Experience is based on student responses to Questions 22 and 23 on the pretest.

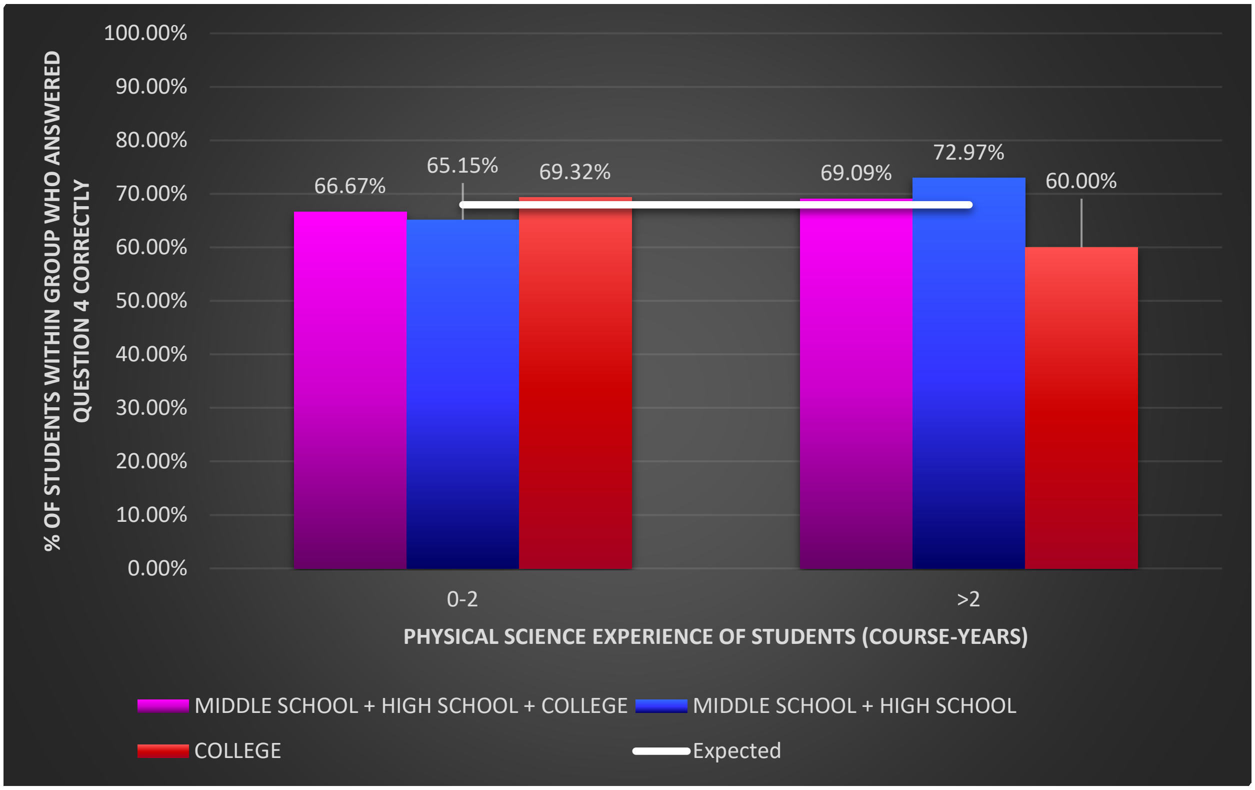

Figure 3. Percentage of students in each experience group who correctly answered “A” on Question 4 of the pretest. Experience is based on student responses to Questions 22 and 23 on the pretest. See the caption for Figure 2 for more information.

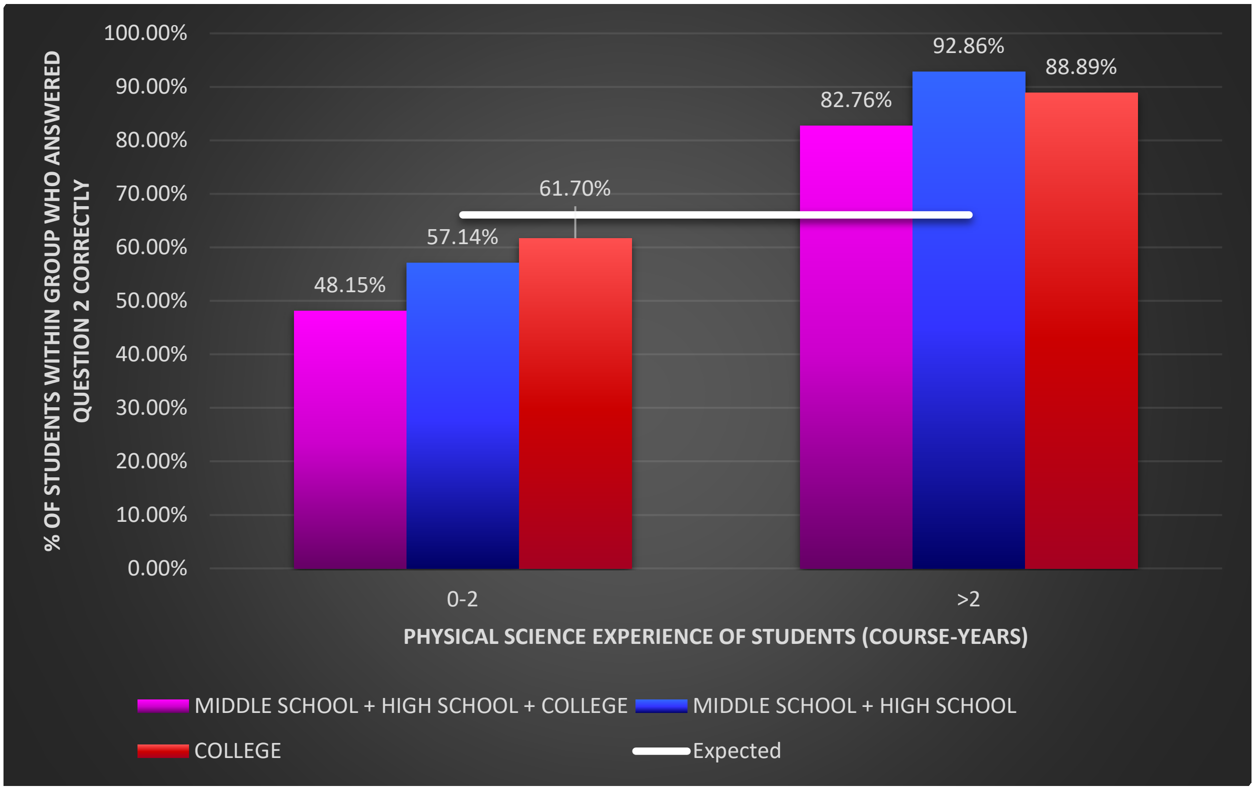

Figure 4. Percentage of students in each experience group who correctly answered “D” on Question 2 of the Post-Exam I survey. Experience is based on student responses to Questions 17 and 18 on the Post-Exam III survey. See the caption for Figure 2 for more information.

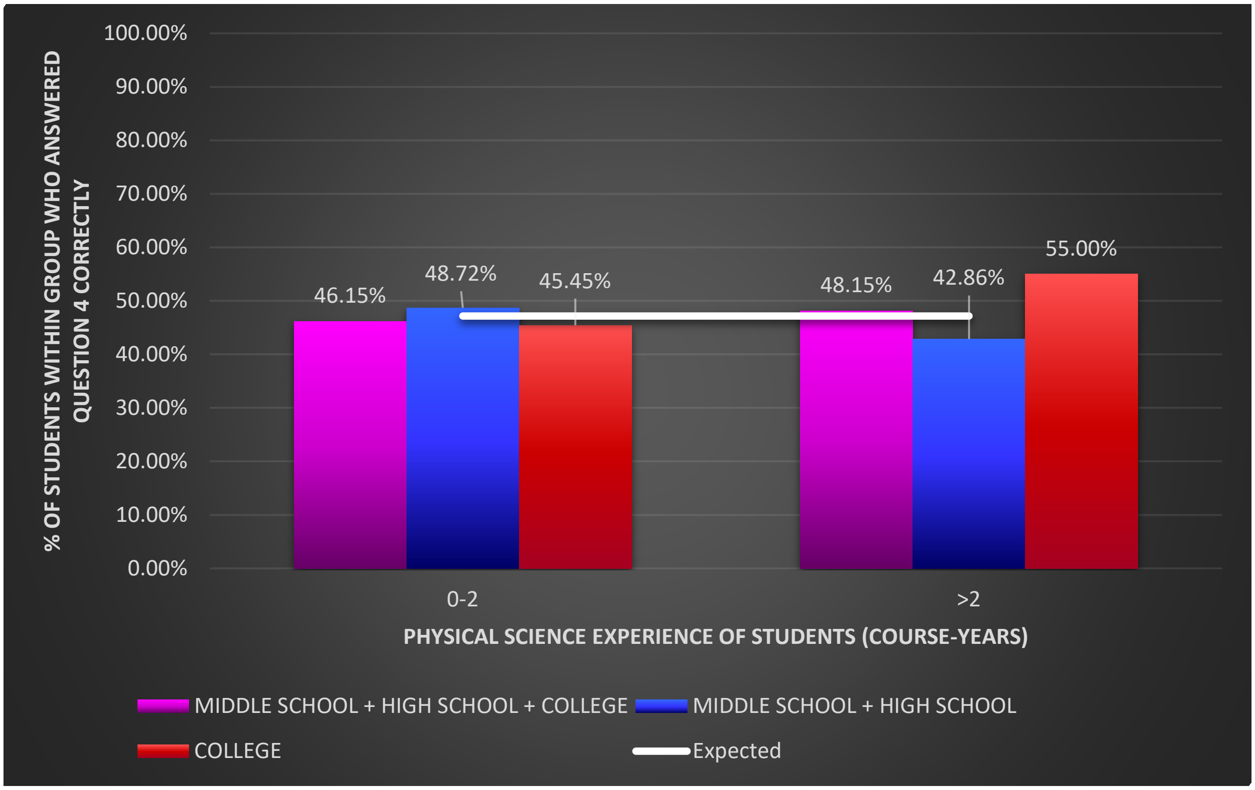

Figure 5. Percentage of students in each experience group who correctly answered “A” on Question 4 of the Post-Exam I survey. Experience is based on student responses to Questions 17 and 18 on the Post-Exam III survey. See the caption for Figure 2 for more information.

For the pretest vs. post-test analysis, most null hypotheses for either test could not be ruled out (see Tables 14-19 for chi-squared values and *p* values). However, the results of the chi-squared test of proportionality for Question 2 when looking at just college experience indicates that the null hypothesis should be rejected, as can be seen on Table 18. This means that I cannot rule out that whether students did or did not change their answers statistically has nothing to do with their previous total or secondary school physical science experience, and that I cannot rule out that whether students did or did not change their answers was independent of their previous college physical science experience, but that I can rule out that students with a low amount of previous college physical science experience changed their answers from incorrect responses on the pretest to correct responses on the Post-Exam I survey at the same frequency as students with a high amount of previous college experience. The conflict between the results of the test of independence and the results of the test of proportionality will be further discussed in Chapter 4.

From the data, I constructed graphs that compare the relative proportions of each experience group who answered a particular question incorrectly on both tests, who answered a particular question correctly on the pretest but not the Post-Exam I survey, who answered a particular question correctly on both tests, and who answered a particular question incorrectly on the pretest but correctly on the Post-Exam I survey (see Figures 6-11).

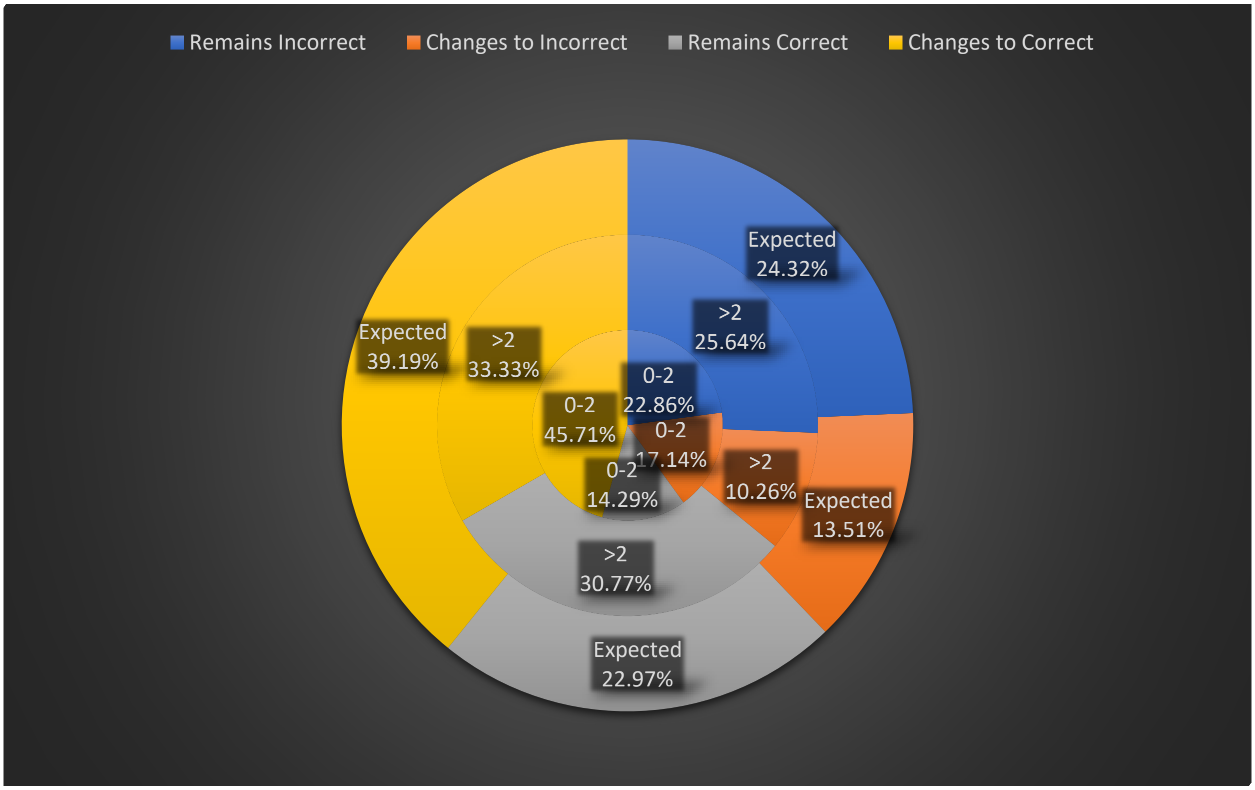

Figure 6. Percentage of students in each experience group who did or did not change their response to Question 2 between the pretest and the Post-Exam I survey, according to their total reported physical science experience (in course-years). The proximal, inner-most circle represents the low experience group; the middle disc represents the high experience group; and the distal, outer-most disc represents the expected proportions for both groups. The blue slice represents students who answered the question incorrectly on both the pretest and the Post-Exam I survey; the orange slice represents students who answered the question correctly on the pretest but incorrectly on the Post-Exam I survey; the grey slice represents students who answered the question incorrectly on the pretest and correctly on the Post-Exam I survey; the yellow slice represents students who answered the question correctly on both surveys. Experience is based on student responses to Questions 22 and 23 on the pretest.

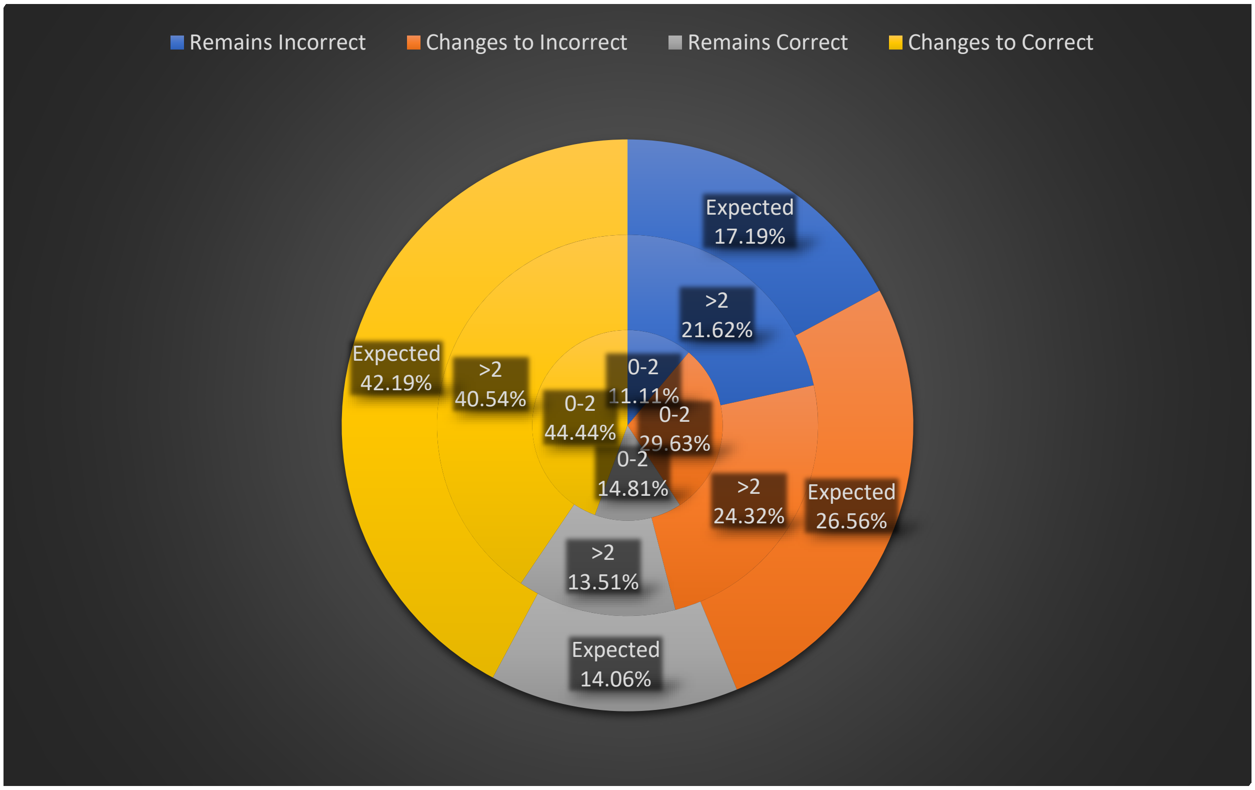

Figure 7. Percentage of students in each experience group who did or did not change their response to Question 4 between the pretest and the Post-Exam I survey, according to their total reported physical science experience (in course-years). Experience is based on student responses to Questions 22 and 23 on the pretest. See the caption for Figure 6 for more information.

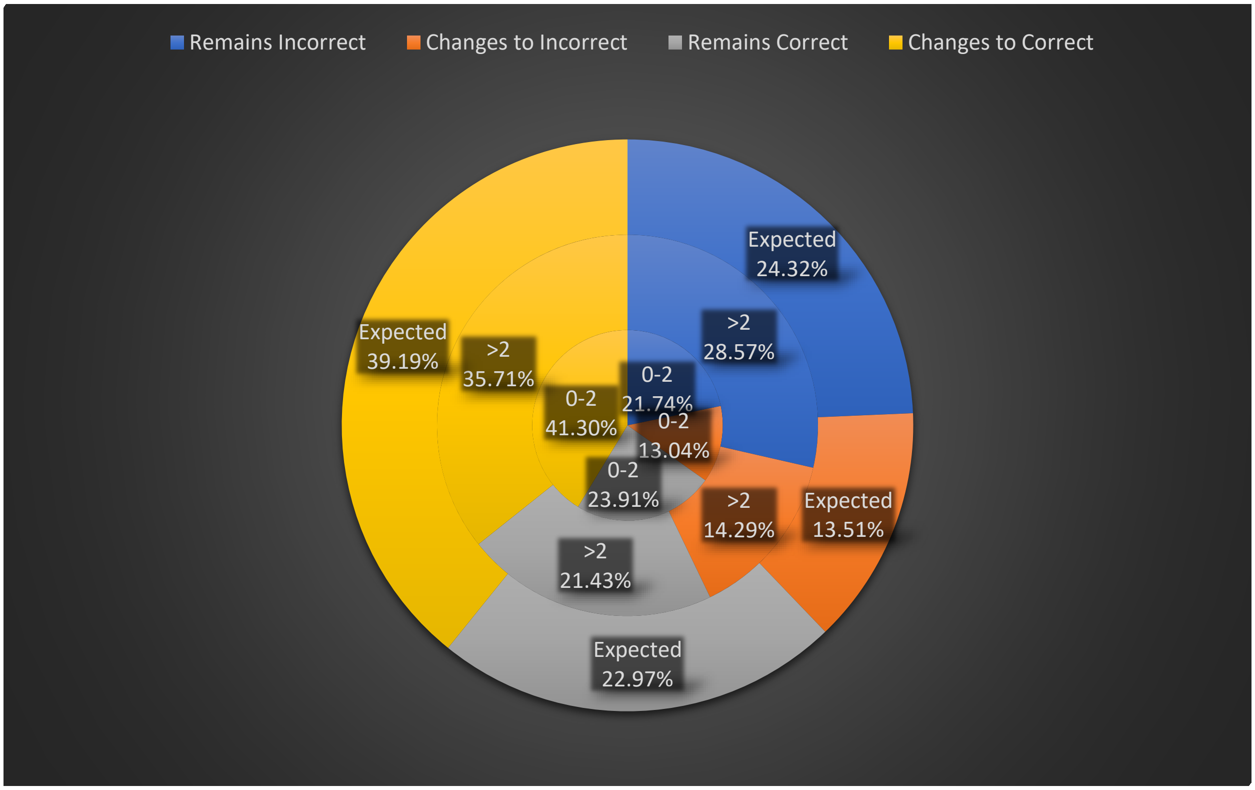

Figure 8. Percentage of students in each experience group who did or did not change their response to Question 2 between the pretest and the Post-Exam I survey, according to their reported secondary school physical science experience (in years). Experience is based on student responses to Question 23 on the pretest. See the caption for Figure 6 for more information.

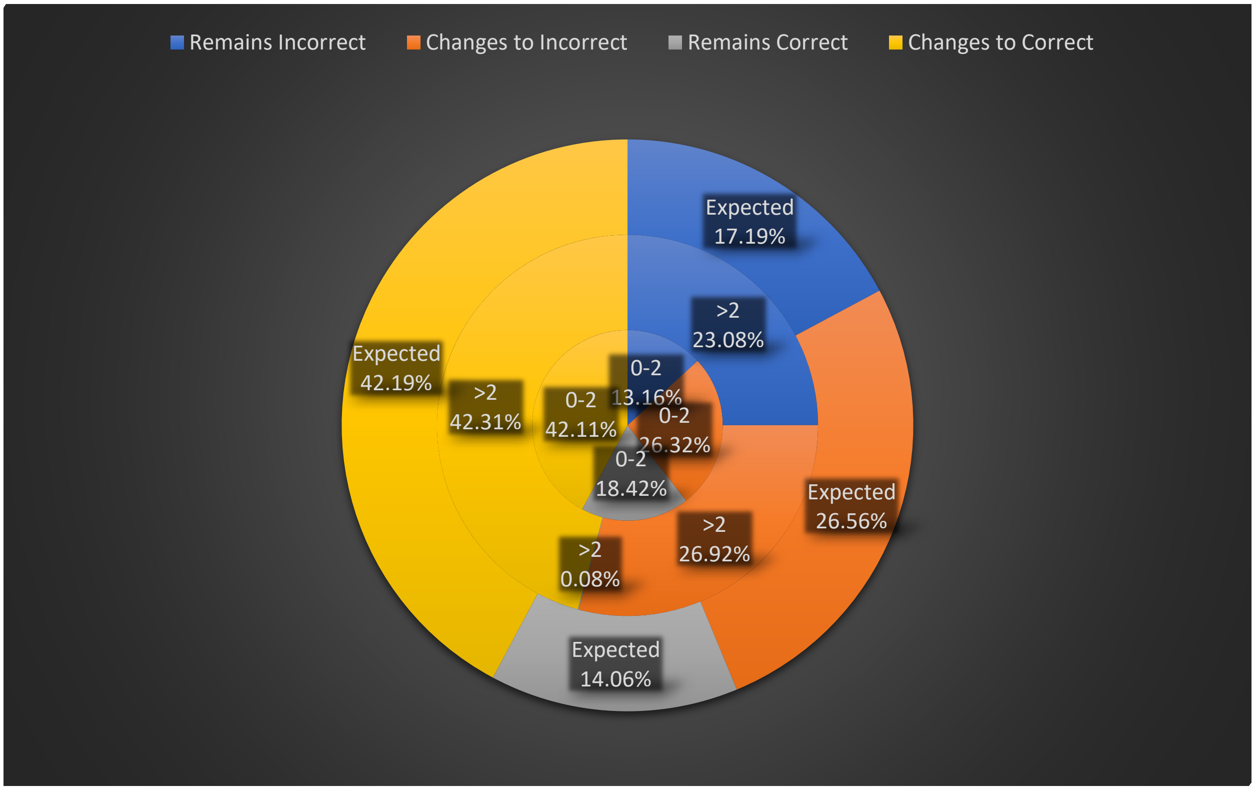

Figure 9. Percentage of students in each experience group who did or did not change their response to Question 4 between the pretest and the Post-Exam I survey, according to their reported secondary school physical science experience (in years). Experience is based on student responses to Question 23 on the pretest. See the caption for Figure 6 for more information.

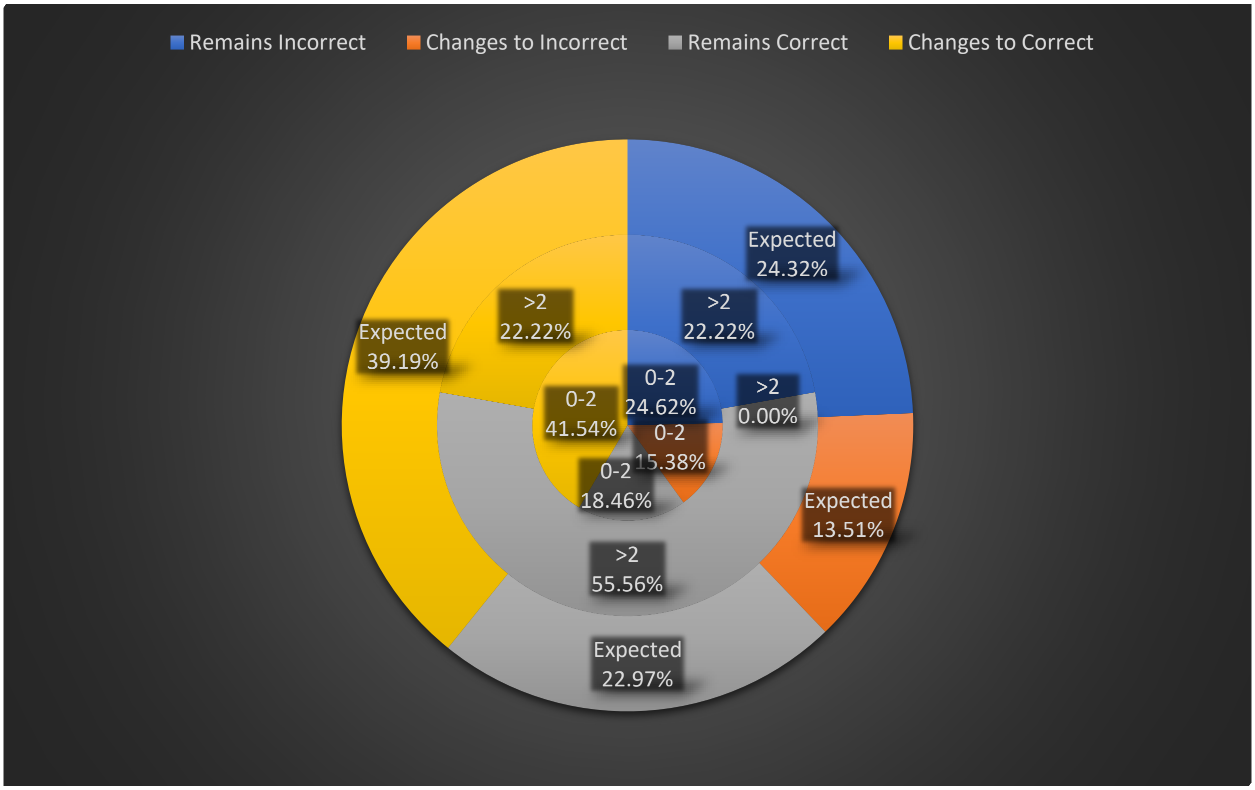

Figure 10. Percentage of students in each experience group who did or did not change their response to Question 2 between the pretest and the Post-Exam I survey, according to their reported physical science experience in college (in courses). Experience is based on student responses to Question 22 on the pretest. See the caption for Figure 6 for more information.

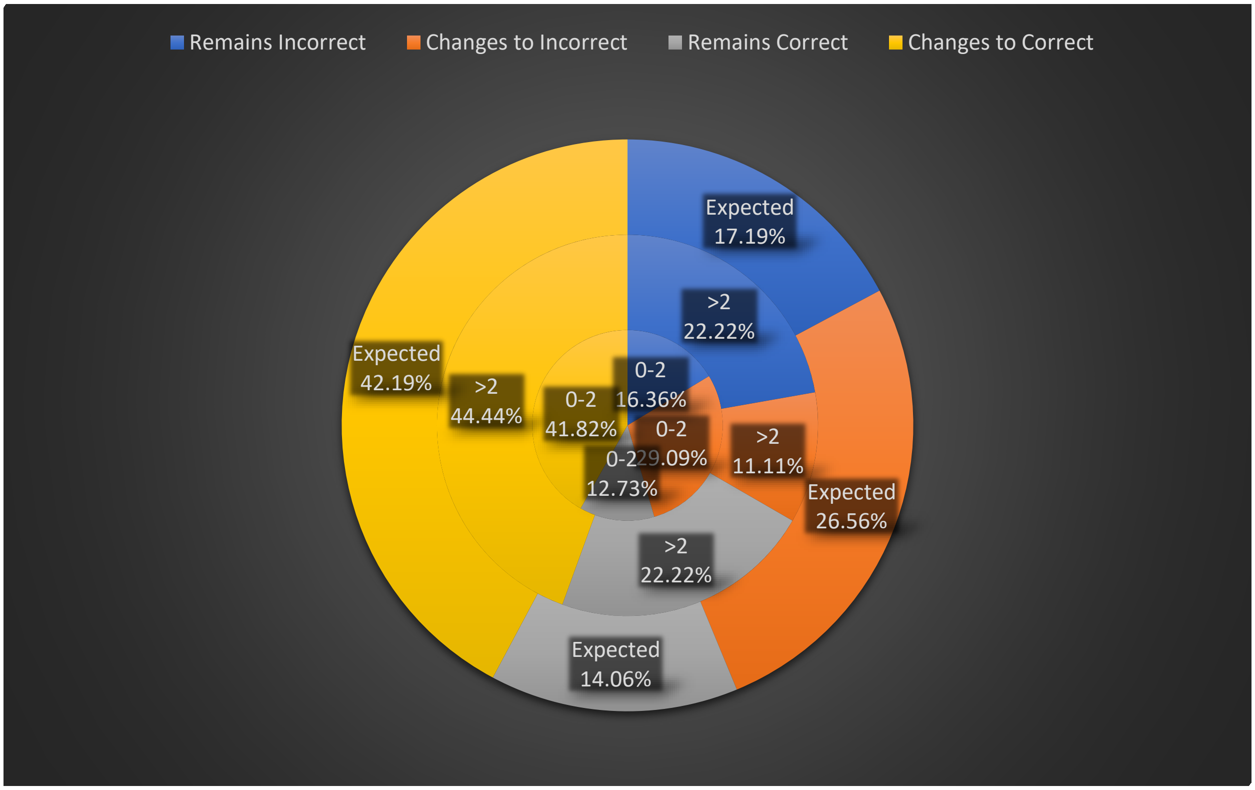

Figure 11. Percentage of students in each experience group who did or did not change their response to Question 4 between the pretest and the Post-Exam I survey, according to their reported physical science experience in college (in courses). Experience is based on student responses to Question 22 on the pretest. See the caption for Figure 6 for more information.

Short-Answer Data

Analyzing Short-Answer Data

Given the issues with the sample size of the multiple-choice data and the inconsistency of the responses to the demographics questions on the pretest and Post-Exam III survey (see Appendix D.1), I chose to investigate the short-answer data using a multi-method analysis of those students who reported the same amount of physical science experience on both the pretest and the Post-Exam III survey. Only 13 students took all the tests and reported consistent amounts of physical science experience on the demographics questions, so to increase the sample size I included all students whose responses to the demographics questions on the Post-Exam III survey did not change the experience group they would be placed in based on their responses to the identical questions on the pretest, regardless of whether the actual numerical responses were consistent. For example, a student who responded "1" on Question 23 of the pretest and "2" on the Post-Exam III survey would be included in the analysis because he or she would remain in the 0-2 secondary school experience group despite their inconsistency; however, a student who responded "2" on the pretest and "3" on the Post-Exam III survey would not be included in the analysis because their experience group is different between the two tests. The final sample size was different for each way of assessing experience: 33 students were included for the analysis that took into account total experience, 34 students were included for the analysis that only considered secondary school experience, and 45 students were included for the analysis that only considered college experience. While those numbers are all certainly larger than 13, they are still rather low. This means that for whatever rich knowledge I have shed light on concerning these students, those results may not hold up when compared against a larger sample size in a different study.

Psychological research such as that by Rosenthal in 1963 has revealed that the beliefs and assumptions of an observer can influence their recorded observations of a subject(s) (known as *experimenter bias* or the *observer-expectancy effect*) (for example, see Rosenthal & Fode, 1963). In order to minimize the observer-expectancy effect on my research results, I blinded myself to the amount of physical science background experience reported by each of the students until the very end of my analysis. Unaware of the reported experience of the 48 students, I went through the students one by one and assessed whether they engaged in transfer and how demonstrably accurate that transfer was concerning four concepts important to answering the question of how temperature and density differences drive plate tectonics: *temperature* ("*T*"), *density* ("*ρ*" or "*rho*"), the *relationship* between temperature and density ("*T-ρ*"), and *vertical motion* ("*V*"). The following is a brief description of what each of those four concepts encompasses:

- **Temperature:** Because the mantle at the asthenosphere-outer core boundary has a relatively fixed temperature, descending mantle material must warm up as it approaches the core. As distance from the core increases, the temperature of solid Earth material decreases.
- **Density:** Denser material will, if allowed, sink underneath less-dense material.
- **Temperature-Density Relationship:** As a material heats up, it expands. This *reduces* the density of that material at any particular point. As a material cools down, it contracts, *increasing* that material's density (no differentiation was made between any student who acknowledged that H_2O uniquely has an inverted temperature-density relationship)
- **Vertical Motion:** Plate tectonics is driven by a convection current, in which some material sinks towards the core while other material rises away from the core.

How much far transfer was demonstrably exhibited by a student in each of these categories was independently coded as either *N* ("no transfer"), *I* ("inaccurate transfer"), *S* ("successful and accurate transfer"), or *A* ("ambiguous"), according to Table 20:

Code	Meaning
N	Seems to exhibit no far transfer of knowledge
I	Previous knowledge is invoked but not demonstrably accurate or relevant
S	Previous knowledge is invoked and is demonstrably accurate
A	Whether or not far transfer occurred is unclear

Table 20. Codes used in the short-answer analysis and what they mean.

For some examples of how student responses were coded, consider the following responses, taken from Appendix D.1:

- *Temperature and density differences between tectonic plates are what cause the movement of the plates. Differing densities of crust which are also at different temperatures allow some plates to move over others while other plates are forces underneath.* (Student #19, pretest)
- *As temperatures change so does density, which will cause circulation. In the case of the earth, denser, cooler materials gravitate towards the core where they are heated, become less dense, and circulate towards the crust, and the cycle continues* (Student #34, Post-Exam I survey)

- *Colder water is more dense than warmer water. When the water warms it becomes less dense and "pushes" against the Earth surface with a lesser amount of pressure. The decrease in pressure caused by the rise in temperature allows the Earth to shift it's plates. I believe this causes natural disasters like tsunamis.* (Student #38, pretest)
- *The higher the temperature and the more dense the earth is in certain areas can cause the earth's plates shift easier.* (Student #98, pretest).
- *It makes the water behave differently which creates different things that move the plates.* (Student #109, pretest)
- *This happens because of the movement of molten rock in the earth. The molten begins near the core where it is heated and we all know that heat rises so the rock is risen further and further away from the core where it also picks up density. The hotter it is the less density it has and the colder it is the more it has. As it reaches a state where the rock can not rise anmore due to weight it sinks back down towards the core. The cycle is again started with the molten rock wwhere it becomes less dense due to the heat. This is how plate tetonics work.* (Student #112, pretest)
- *I'd assume that plates with looser and hotter substances around them would be less stable and more likely to shift than plates with denser and colder substances around them.* (Student #122, pretest)

Student #19's pretest response is lacking, but there does seem to be some transfer happening. The student brings up temperature and density, and mentions that differences in temperature and density cause plate motion. However, because no mention is made of the difference in the motion of warm material versus cold material or denser material versus less-dense material, I cannot conclude with certainty that this transfer is accurate. It is also completely unclear whether or not the student recognizes an actual connection between temperature and density, and no mention is made of *vertical* motion. This response received the codes {I, I, A, N} for T, ρ, T-ρ, and V respectively.

Student #34's Post-Exam I survey response is a classic example of a completely accurate response with a lot of successful transfer of physical science concepts on display. Responses like Student #34's were coded as {S, S, S, S}.

Student #38's pretest response, which is highly erroneous, makes no reference to vertical motion and has very little to say about temperature. However, Student #38 accurately mentions that a substance's density decreases as its temperature increases, and while what they had to say about density itself was not demonstrably accurate, it was clear that the student was engaging in some form of transfer. This responses was coded {N, I, S, N}.

Student #98's pretest response was highly ambiguous. They make reference to hotter materials and denser materials, but it is unclear whether they (erroneously) consider a substance's density to increase with temperature, nor is it clear if they are transferring any relevant information about temperature and density on their own. They also do not indicate the direction of plate "shifting" that they are talking about. This response would be coded as {A, A, A, N}.

Student #109's pretest response is completely inaccurate and makes no reference to any of the concept categories of interest to this study. It was thus coded as {N, N, N, N}.

Student #112's pretest response falsely suggests that the mantle is composed of liquid rock, but the actual accuracy of a student's response is not important to my study, just whether transfer of particular physical science concepts occurred, so the inaccuracy did not affect the coding process. Because the student's response shows clear transfer of the temperature, *T*-ρ, and vertical motion concept categories, the response received the "S" code in all of those categories. However, while the student brings density's relationship to temperature up, no indication is made that density has anything to do with mantle motion (the student instead inaccurately brings up the idea of *weight*, a physical force affected by mass), so the student received an "A" for the density concept category. Overall, the codes for Student #112's response was {S, A, S, S}.

Student #122's pretest response only makes passing reference to temperature and density and no mention of the vertical motion of the mantle; they do mention substances that are *denser* and *colder*, but because these two descriptors are not explicitly linked by the student, it is difficult to tell if the student understands that the temperature of a substance is directly related to its density. Therefore, I would code Student #122's reponse as {N, N, A, N}.

In the course of coding the responses, I came across various situations that challenged my coding scheme. The relationship between temperature and density was very rarely explicitly stated. This seems to be due to a limitation of the question, which did not explicitly request that students clarify that thermal expansion in high temperatures results in a reduction of density. To resolve this issue, the judgment call was made to allow phrases that appear to make an implicit connection between temperature and density, such as *…the less dense, hotter material…* to qualify for an "S." Some students, as will be discussed later, explicitly brought up (mantle) convection, without indicating explicitly that the convection was driven by temperature and density. While I argue that use of the term *convection*, an uncommon word amongst the general public, is a sign that transfer of physical science background knowledge is surely occurring, it does not necessarily imply an understanding of the underlying dynamics; therefore, a response that brought up convection but did not discuss or insufficiently discussed temperature and density only received an “S” for the concept of vertical motion. In the analysis, students occasionally brought up physical science concepts that were demonstrably accurate, a clear sign of far transfer, but were not demonstrably relevant to the topic of plate tectonics. An example of such a concept was the idea that objects with different composition but that are otherwise identical have different densities. This concept is still somewhat relevant to the density concept being measured, but density differences due to composition have not been found to contribute meaningfully to mantle convection beyond the fact that continental lithospheric material is not dense enough to sink into the asthenosphere. I thus made the decision to code such a response as “I.”

My background in coding is limited. To combat any potential biases in how I coded the responses, I went through and coded the responses three separate times, each at a separate sitting. All responses were coded in each session, and I took great care to blind myself of codes that I previously assigned each response in the previous session(s). During the first two sessions, I coded responses starting with those of the student with ID number 1 and worked my way in increasing numerical order, ending with the student with ID number 122; for the third coding session, I started with the pretest and Post-Exam I survey responses of the student with ID number 122 and worked my way in decreasing numerical order, ending with the student with ID number 1. During each session, for each particular student I read their pretest response to Question 6, coded the response with the appropriate letters (one for each category), read their Post-Exam I survey response to Question 7, coded that response, and jotted down any qualitative remarks I had about the student’s responses and how they changed between the pretest and Post-Exam I survey before moving on to the next student’s pretest response and starting the process over. While the quantitative part of the analysis only used the codes from the third coding session, the codes and remarks from all three sessions are viewable in Appendix D.1. As previously mentioned, while the amount of physical science background experience each student had may be viewed in Appendix D.1, this information was hidden from me during the entire coding process and was never revealed to me until the end of the quantitative portion of the analysis.

After every response was coded thrice, a raw frequency table was drawn up for each concept category based on the codes from the third coding session (see Appendix D.2-D.5), charting the frequency of each code on the pretest and on the Post-Exam I survey organized by reported physical science background experience, similar to those that I used in the multiple-choice data analyses. Appendix D.6-D.29 summarize Appendix D.2-D.5 by compressing the various amounts of experience into our two familiar experience groups and calculating a mode and variation ratio for each. Due to the categorical nature of the data, I decided that the best way to measure each experience group's central tendency would be to identify that group's statistical mode code. In order to measure how variable each group was in their degree of transfer, I also calculated each mode's associated variation ratio (the proportion of responses that were given a different code than the mode). The formula used to calculate the observed variation ratio was:

$$v = 1 - \frac{f_{mode}}{n}$$

In the above formula, v is the variation ratio, f_{mode} is the frequency of the observed mode of the particular experience group in question, and n is the sample size of the experience group. The calculated expected values for the mode and variation ratio based on the distribution of the data were identical to the respective observed values for the total sample.

From Appendix D.6-D.29, I produced three tables and 24 graphs. Tables 21-23 show the mode code(s) for each concept category depending on the experience group (0-2 and >2) and which test the response was for (pretest or Post-Exam I survey, the latter labeled as *post-test*). The graphs drawn up for each concept category (see Figures 12-35) show the percentage of students within an experience group whose responses received a particular code. A line on each column indicates the corresponding expected percent frequency of each code on the pretest. Expected values are based on the expected statistical distribution of the codes on the pretest and were calculated similar to how one would calculate them for a chi-squared statistical test. Two graphs each were made for total experience, secondary school experience, and college experience: one for the pretest responses and one for the Post-Exam I survey responses, for a total of 24 graphs.

The decision to compare observed Post-Exam I survey percent frequencies, modes, and variation ratios to expected *pretest* values rather than Post-Exam I values was made because doing so was deemed to better investigate whether having more experience promoted more growth over the course of the semester.

Once the quantitative portion of the short-answer analysis was complete, I took one final look at the responses before removing the blinds over each student's background experience. I then proceeded with the qualitative portion of the analysis. My observations were based on the coding from the three sessions (particularly those from the second and third sessions, as they were less biased by novelty), the remarks that I made from all three coding sessions, general trends I saw while arranging the data in particular orders (i.e. by ascending total experience), and the results of the quantitative portion of the analysis.

Mode Degree of Transfer on Question 6 of the Pretest and Question 7 of the Post-Exam I Survey According to Total Reported Previous Physical Science Experience										
PRETEST										
N (# of students)		33	**TEMPERATURE**		**DENSITY**		***T*-*ρ* RELATIONSHIP**		**VERTICAL MOTION**	
	Bin	***n*** (# of students)	**Mode**	**Variation Ratio**	**Mode**	**Variation Ratio**	**Mode**	**Variation Ratio**	**Mode**	**Variation Ratio**
Total Experience	**0-2**	19	**N**	0.1053	**N**	0.0526	**N**	0.1053	**N**	0.1053
	>2	14	**N**	0.2857	**N**	0.4286	**N**	0.6429	**N**	0.1429
	Total	33	**N**	0.1818	**N**	0.2121	**N**	0.3333	**N**	0.1212
POST-EXAM I SURVEY										
N (# of students)		33	**TEMPERATURE**		**DENSITY**		***T*-*ρ* RELATIONSHIP**		**VERTICAL MOTION**	
	Bin	***n*** (# of students)	**Mode**	**Variation Ratio**	**Mode**	**Variation Ratio**	**Mode**	**Variation Ratio**	**Mode**	**Variation Ratio**
Total Experience	**0-2**	19	**N**	0.3158	**N**	0.5263	**N**	0.2632	**N**	0.2105
	>2	14	**S**	0.5000	**S**	0.6429	**N**	0.5714	**S**	0.4286
	Total	33	**N**	0.5455	**N**	0.6061	**N**	0.3939	**N**	0.4242

Table 21. Mode degree to which students transferred previous knowledge about physical science when answering Question 6 on the pretest and Question 7 on the Post-Exam I survey, broken down by reported total experience and by concept category. Variation ratios are provided for each mode to give a sense of how variable each group was in their degree of transfer. The modes and variation ratios for the total sample are equal to the expected values for their column. Total experience (in course-years) was calculated based on student responses to Questions 22 and 23 on the pretest. See the caption for Table 2 for more information.

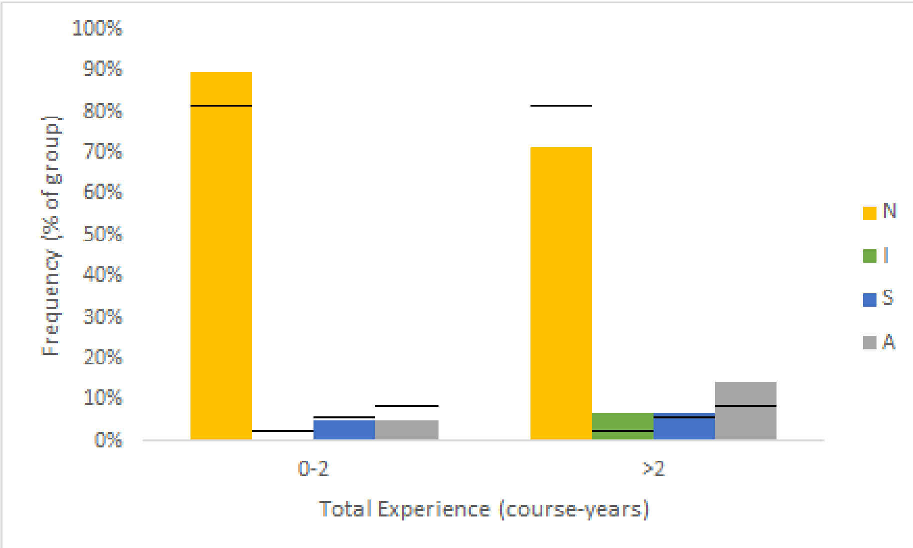

Figure 12. Degree of transfer of knowledge about temperature by total experience group on Question 6 of the pretest. The orange bar represents the percentage of students within an experience group whose response was coded as "N" (see Table 20 for information on the coding scheme); the green bar represents the percentage of students within an experience group whose response was coded as "I;" the blue bar represents the percentage of students within an experience group whose response was coded as "S;" the grey bar represents the percentage of students within an experience group whose response was coded as "A." The lines represent the expected percent frequencies for responses to Question 6 on the pretest (see Appendix D.7 or relevant appendix on sheet). Experience is based on student responses to Questions 22 and 23 on the pretest.

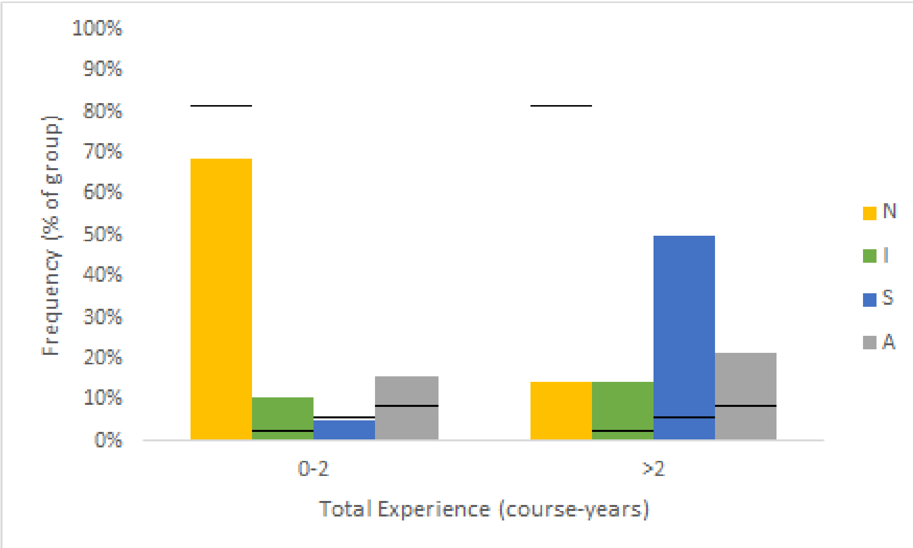

Figure 13. Degree of transfer of knowledge about temperature by total experience group on Question 7 of the Post-Exam I survey. Experience is based on student responses to Questions 17 and 18 on the Post-Exam III survey. Expected values are based on pretest data (see Appendix D.7 or relevant appendix on sheet). See the caption for Figure 12 for more information.

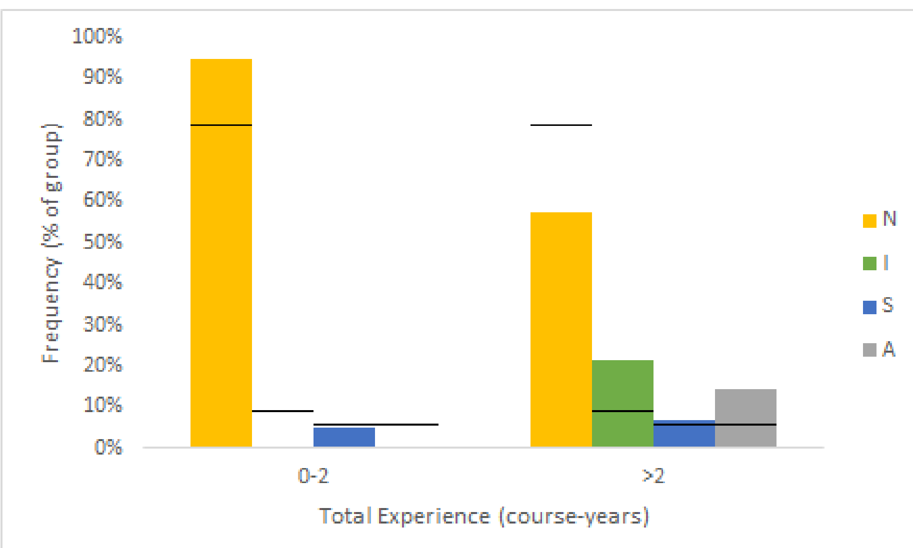

Figure 14. Degree of transfer of knowledge about density by total experience group on Question 6 of the pretest. Experience is based on student responses to Questions 22 and 23 on the pretest. See the caption for Figure 12 for more information.

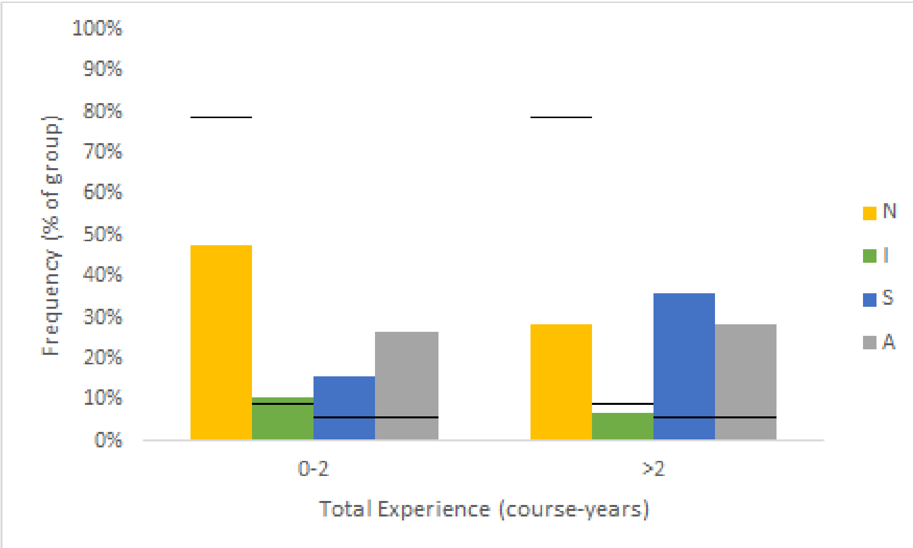

Figure 15. Degree of transfer of knowledge about density by total experience group on Question 7 of the Post-Exam I survey. Experience is based on student responses to Questions 17 and 18 on the Post-Exam III survey. Expected values are based on pretest data. See the caption for Figure 12 for more information.

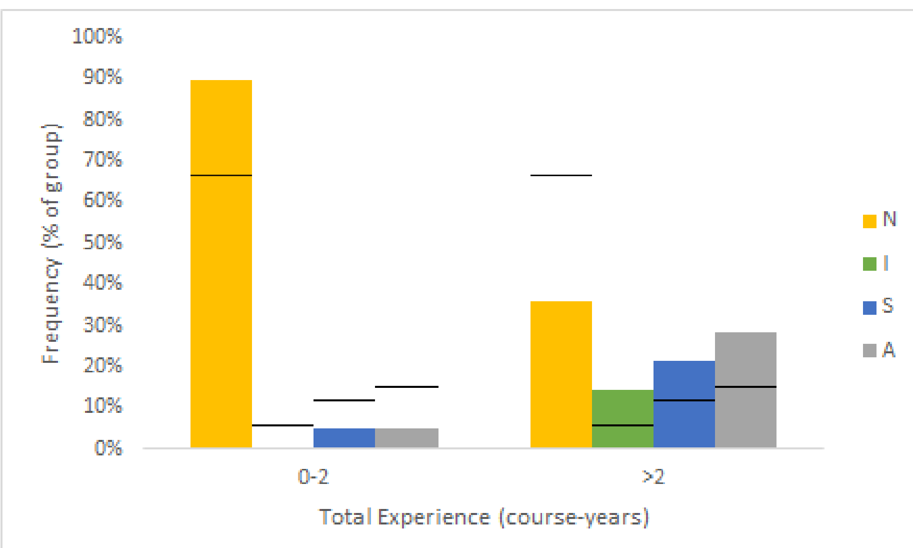

Figure 16. Degree of transfer of knowledge about the relationship between temperature and density by total experience group on Question 6 of the pretest. Experience is based on student responses to Questions 22 and 23 on the pretest. See the caption for Figure 12 for more information.

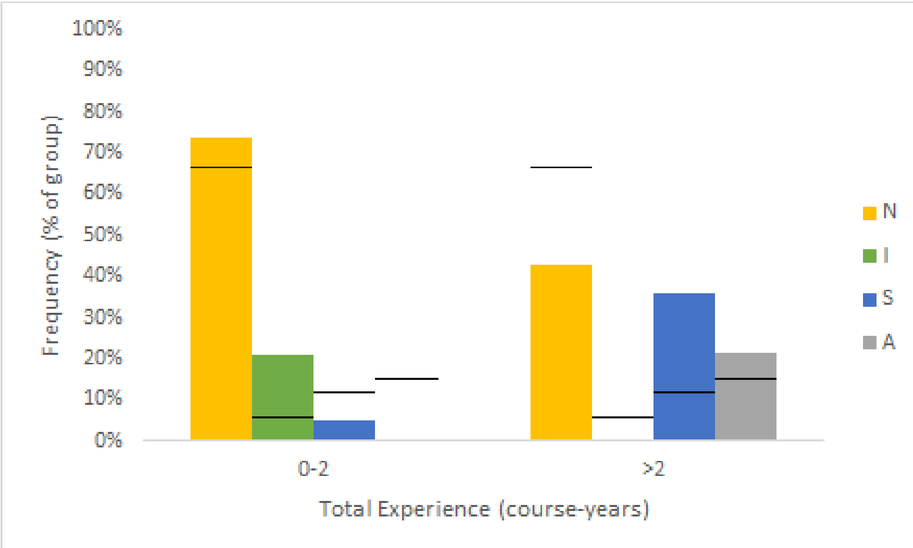

Figure 17. Degree of transfer of knowledge about the relationship between temperature and density by total experience group on Question 7 of the Post-Exam I survey. Experience is based on student responses to Questions 17 and 18 on the Post-Exam III survey. Expected values are based on pretest data. See the caption for Figure 12 for more information.

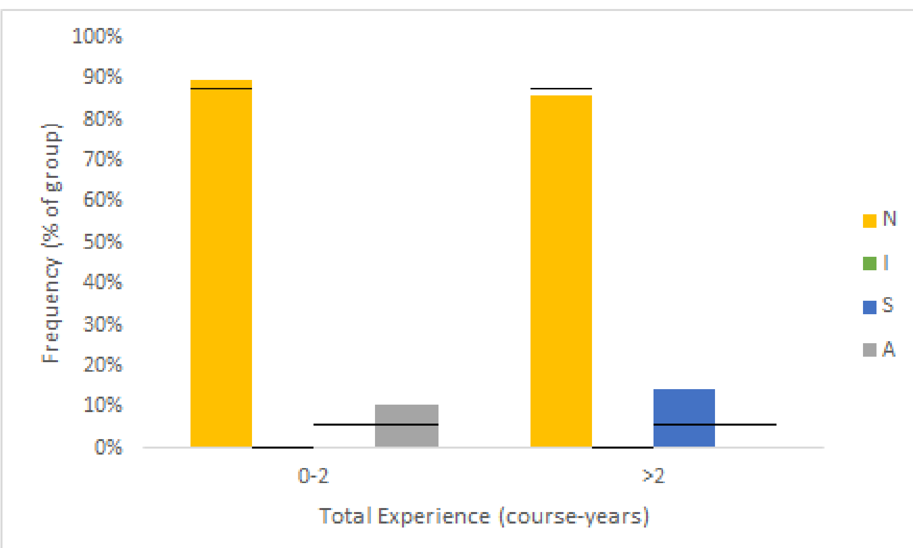

Figure 18. Degree of transfer of knowledge about vertical motion by total experience group on Question 6 of the pretest. Experience is based on student responses to Questions 22 and 23 on the pretest. See the caption for Figure 12 for more information.

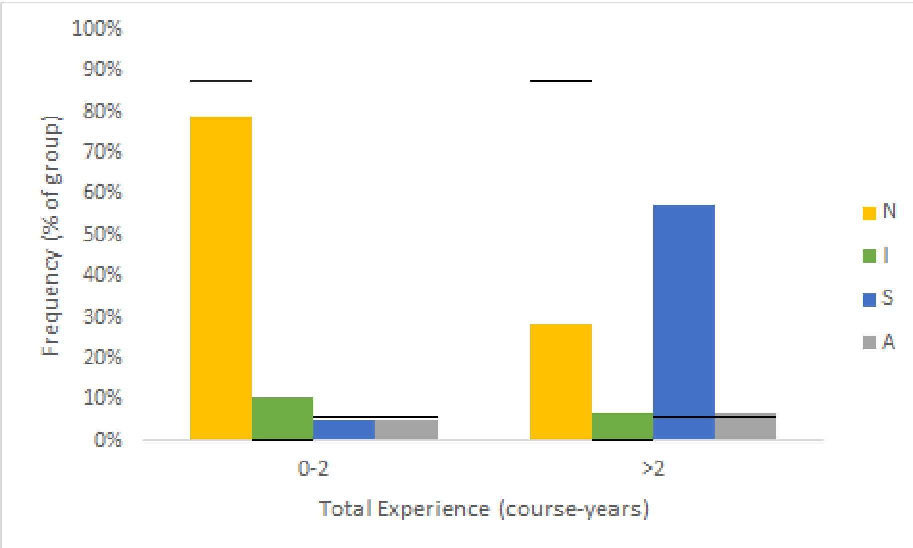

Figure 19. Degree of transfer of knowledge about vertical motion by total experience group on Question 7 of the Post-Exam I survey. Experience is based on student responses to Questions 17 and 18 on the Post-Exam III survey. Expected values are based on pretest data. See the caption for Figure 12 for more information.

Mode Degree of Transfer on Question 6 of the Pretest and Question 7 of the Post-Exam I Survey According to Reported Previous Physical Science Experience in Secondary School										
PRETEST										
N (# of students)		34	**TEMPERATURE**		**DENSITY**		***T*-*ρ* RELATIONSHIP**		**VERTICAL MOTION**	
	Bin	***n*** (# of students)	**Mode**	**Variation Ratio**	**Mode**	**Variation Ratio**	**Mode**	**Variation Ratio**	**Mode**	**Variation Ratio**
Secondary School Experience	**0-2**	29	**N**	0.1379	**N**	0.1379	**N**	0.2069	**N**	0.1724
	>2	5	**N**	0.4000	**N**	0.4000	**S**	0.6000	**N**	0.0000
	Total	34	**N**	0.1765	**N**	0.1765	**N**	0.2941	**N**	0.1471
POST-EXAM I SURVEY										
N (# of students)		34	**TEMPERATURE**		**DENSITY**		***T*-*ρ* RELATIONSHIP**		**VERTICAL MOTION**	
	Bin	***n*** (# of students)	**Mode**	**Variation Ratio**	**Mode**	**Variation Ratio**	**Mode**	**Variation Ratio**	**Mode**	**Variation Ratio**
Secondary School Experience	**0-2**	29	**N**	0.4138	**N**	0.5172	**N**	0.3448	**N**	0.2414
	>2	5	**S**	0.4000	**SA**	0.2000	**S**	0.4000	**S**	0.2000
	Total	34	**N**	0.5000	**N**	0.5882	**N**	0.4118	**N**	0.3529

Table 22. Mode degree to which students transferred previous knowledge about physical science when answering Question 6 on the pretest and Question 7 on the Post-Exam I survey, broken down by reported secondary school experience and by concept category. Secondary school experience (in years) was based on student responses to Question 23 on the pretest. See the captions for Table 2 and Table 21 for more information.

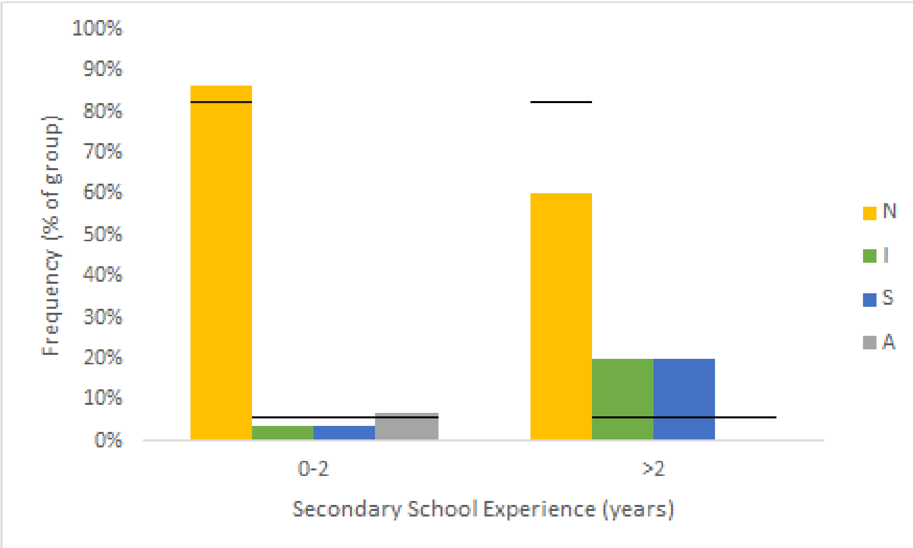

Figure 20. Degree of transfer of knowledge about temperature by secondary school experience group on Question 6 of the pretest. Experience is based on student responses to Question 23 on the pretest. See the caption for Figure 12 for more information.

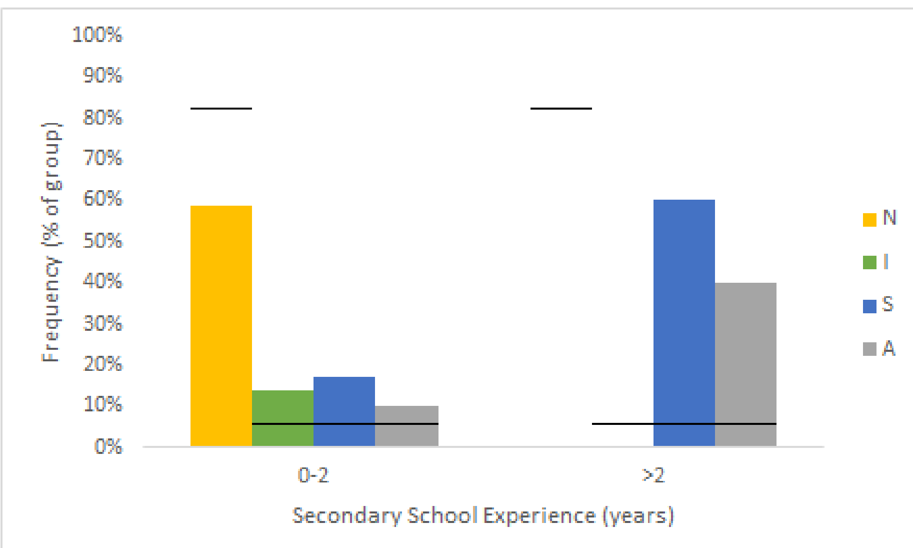

Figure 21. Degree of transfer of knowledge about temperature by secondary school experience group on Question 7 of the Post-Exam I survey. Experience is based on student responses to Question 18 on the Post-Exam III survey. Expected values are based on pretest data. See the caption for Figure 12 for more information.

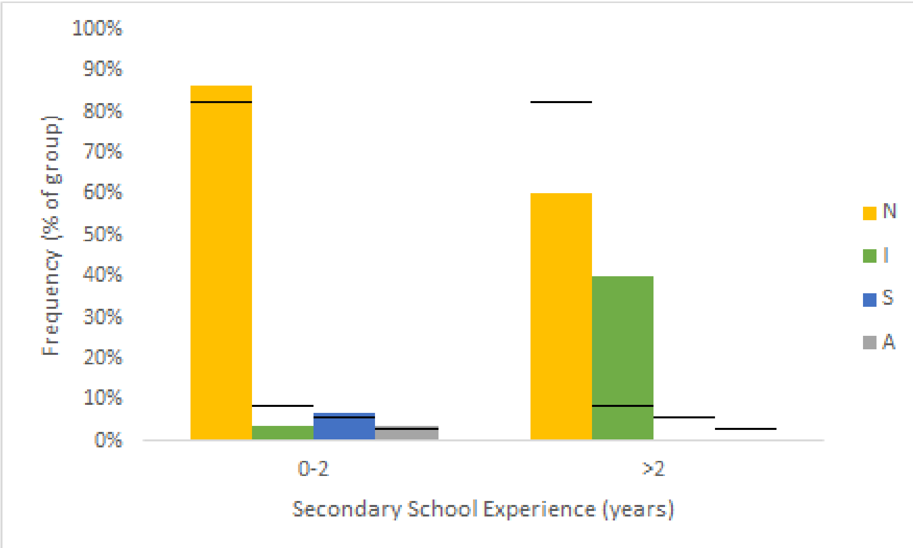

Figure 22. Degree of transfer of knowledge about density by secondary school experience group on Question 6 of the pretest. Experience is based on student responses to Question 23 on the pretest. See the caption for Figure 12 for more information.

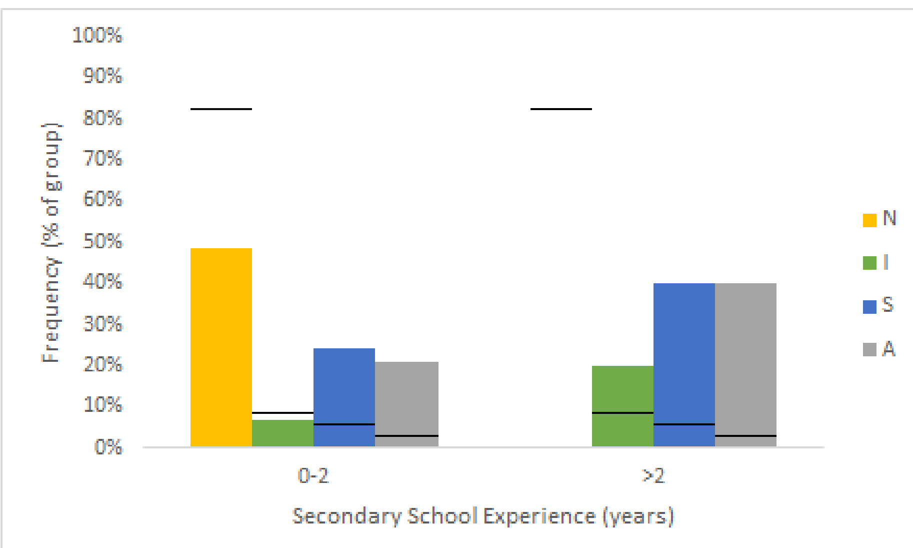

Figure 23. Degree of transfer of knowledge about density by secondary school experience group on Question 7 of the Post-Exam I survey. Experience is based on student responses to Question 18 on the Post-Exam III survey. Expected values are based on pretest data. See the caption for Figure 12 for more information.

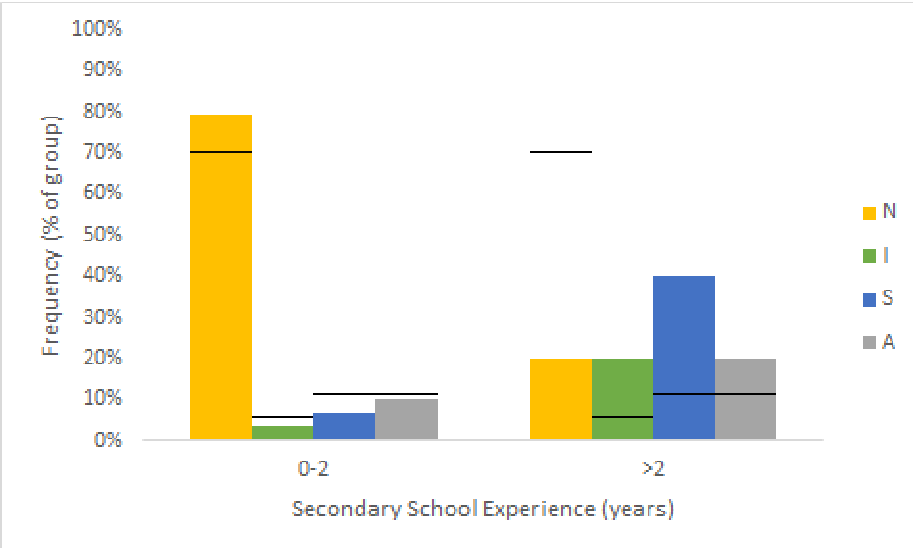

Figure 24. Degree of transfer of knowledge about the relationship between temperature and density by secondary school experience group on Question 6 of the pretest. Experience is based on student responses to Question 23 on the pretest. See the caption for Figure 12 for more information.

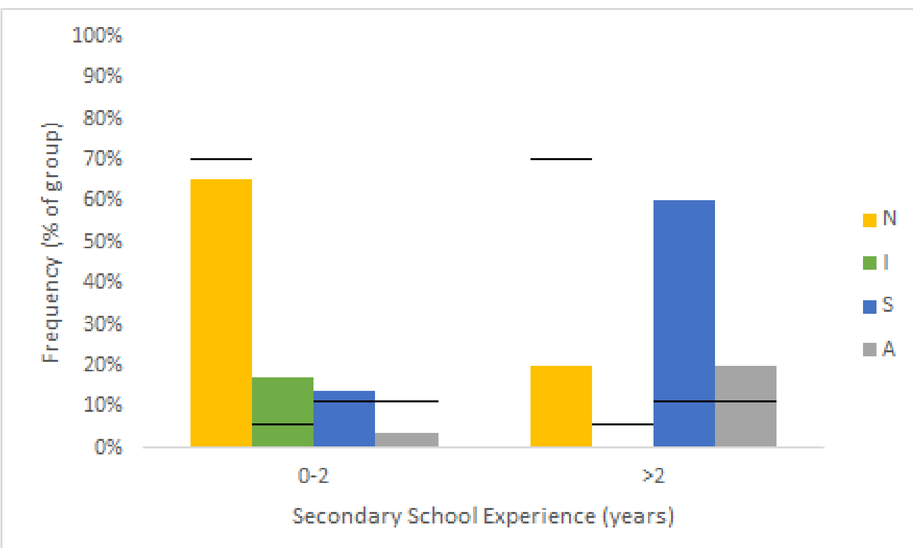

Figure 25. Degree of transfer of knowledge about the relationship between temperature and density by secondary school experience group on Question 7 of the Post-Exam I survey. Experience is based on student responses to Question 18 on the Post-Exam III survey. Expected values are based on pretest data. See the caption for Figure 12 for more information.

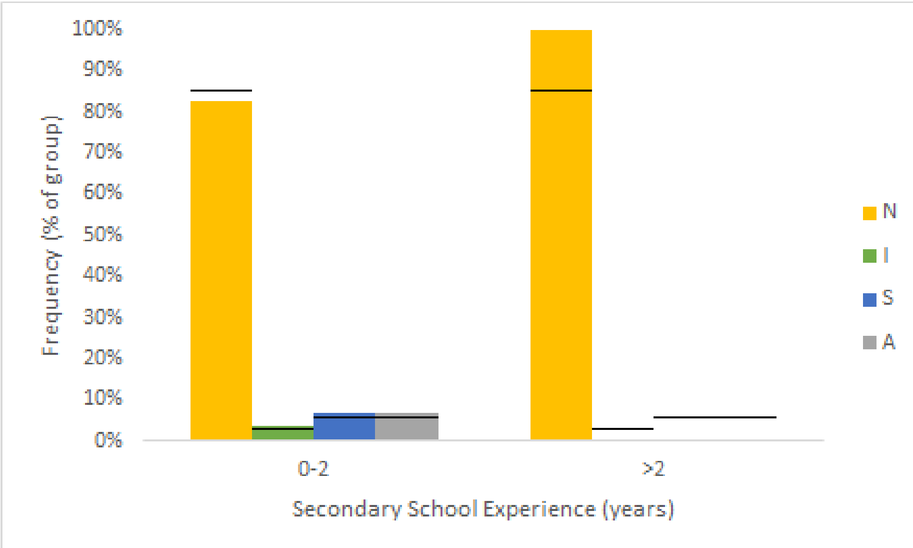

Figure 26. Degree of transfer of knowledge about vertical motion by secondary school experience group on Question 6 of the pretest. Experience is based on student responses to Question 23 on the pretest. See the caption for Figure 12 for more information.

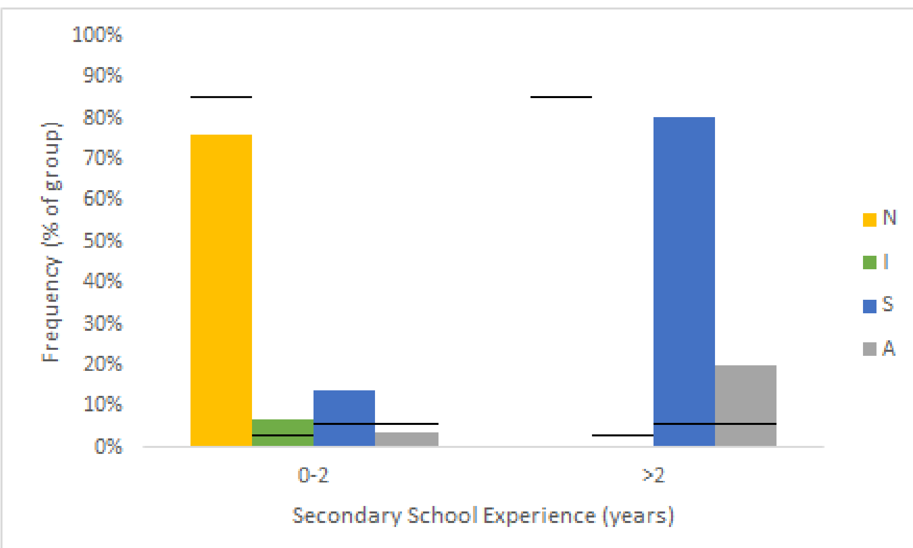

Figure 27. Degree of transfer of knowledge about vertical motion by secondary school experience group on Question 7 of the Post-Exam I survey. Experience is based on student responses to Question 18 on the Post-Exam III survey. Expected values are based on pretest data. See the caption for Figure 12 for more information.

Mode Degree of Transfer on Question 6 of the Pretest and Question 7 of the Post-Exam I Survey According to Reported Previous Physical Science Experience in College										
PRETEST										
N (# of students)		45	**TEMPERATURE**		**DENSITY**		***T*-*ρ* RELATIONSHIP**		**VERTICAL MOTION**	
	Bin	***n*** (# of students)	**Mode**	**Variation Ratio**	**Mode**	**Variation Ratio**	**Mode**	**Variation Ratio**	**Mode**	**Variation Ratio**
College Experience	**0-2**	39	**N**	0.2564	**N**	0.1538	**N**	0.2564	**N**	0.1538
	>2	6	**N**	0.3333	**I**	0.5000	**N**	0.5000	**N**	0.0000
	Total	45	**N**	0.2667	**N**	0.2222	**N**	0.2889	**N**	0.1333
POST-EXAM I SURVEY										
N (# of students)		45	**TEMPERATURE**		**DENSITY**		***T*-*ρ* RELATIONSHIP**		**VERTICAL MOTION**	
	Bin	***n*** (# of students)	**Mode**	**Variation Ratio**	**Mode**	**Variation Ratio**	**Mode**	**Variation Ratio**	**Mode**	**Variation Ratio**
College Experience	**0-2**	39	**N**	0.3590	**N**	0.4872	**N**	0.3846	**N**	0.3077
	>2	6	**S**	0.5000	**NSA**	0.0000	**S**	0.5000	**S**	0.5000
	Total	45	**N**	0.4222	**N**	0.5111	**N**	0.4222	**N**	0.3556

Table 23. Mode degree to which students transferred previous knowledge about physical science when answering Question 6 on the pretest and Question 7 on the Post-Exam I survey, broken down by reported college experience and by concept category. The modes and variation ratios for the total sample are equal to the expected values for their column. College experience (in courses) was based on student responses to Question 22 on the pretest. See the captions for Table 2 and Table 21 for more information.

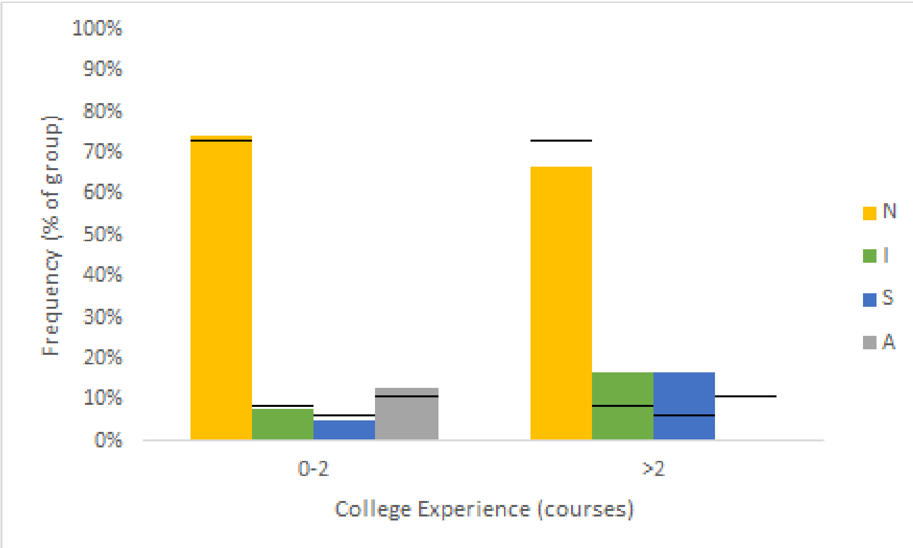

Figure 28. Degree of transfer of knowledge about temperature by college experience group on Question 6 of the pretest. Experience is based on student responses to Question 22 on the pretest. See the caption for Figure 12 for more information.

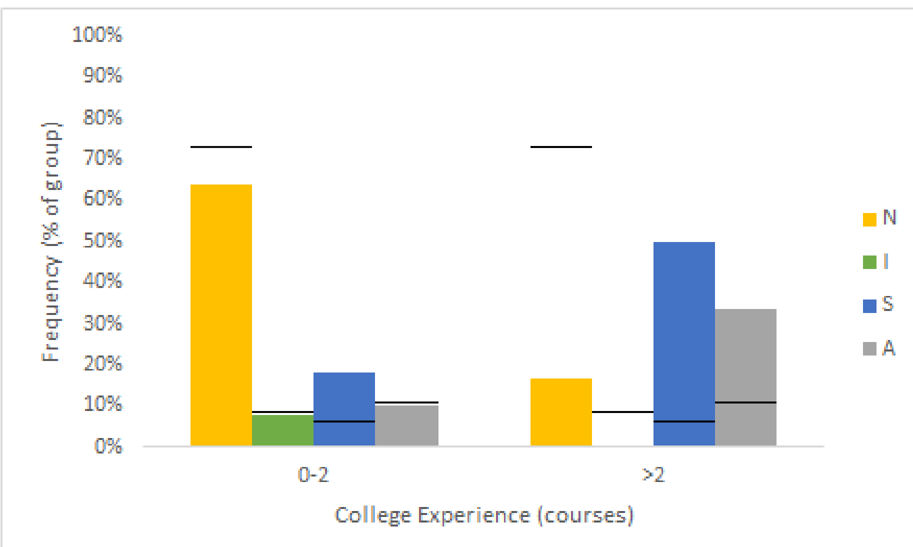

Figure 29. Degree of transfer of knowledge about temperature by college experience group on Question 7 of the Post-Exam I survey. Experience is based on student responses to Question 17 on the Post-Exam III survey. Expected values are based on pretest data. See the caption for Figure 12 for more information.

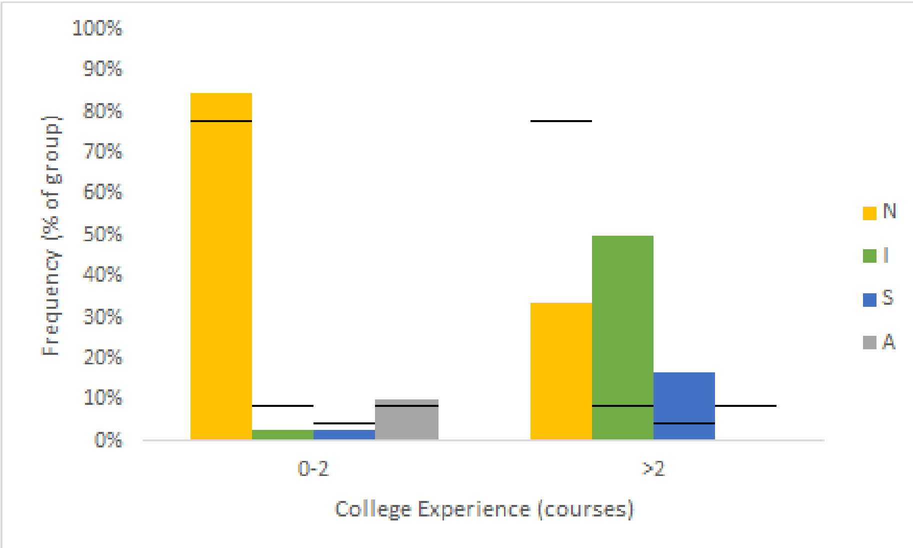

Figure 30. Degree of transfer of knowledge about density by college experience group on Question 6 of the pretest. Experience is based on student responses to Question 22 on the pretest. See the caption for Figure 12 for more information.

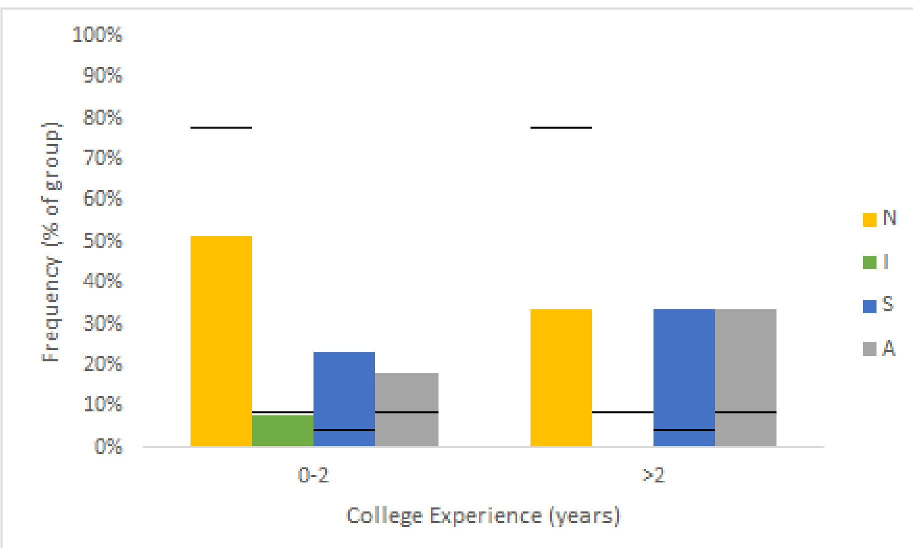

Figure 31. Degree of transfer of knowledge about density by college experience group on Question 7 of the Post-Exam I survey. Experience is based on student responses to Question 17 on the Post-Exam III survey. Expected values are based on pretest data. See the caption for Figure 12 for more information.

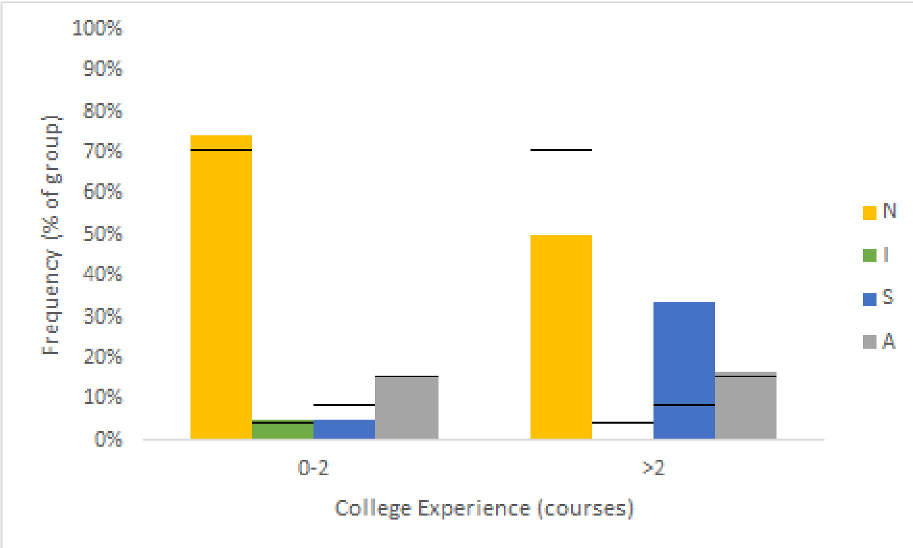

Figure 32. Degree of transfer of knowledge about the relationship between temperature and density by college experience group on Question 6 of the pretest. Experience is based on student responses to Question 22 on the pretest. See the caption for Figure 12 for more information.

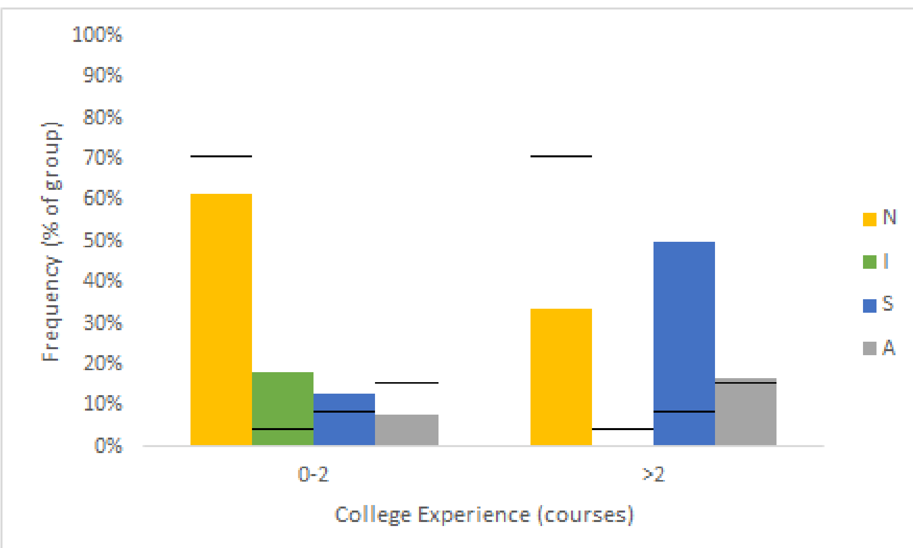

Figure 33. Degree of transfer of knowledge about the relationship between temperature and density by college experience group on Question 7 of the Post-Exam I survey. Experience is based on student responses to Question 17 on the Post-Exam III survey. Expected values are based on pretest data. See the caption for Figure 12 for more information.

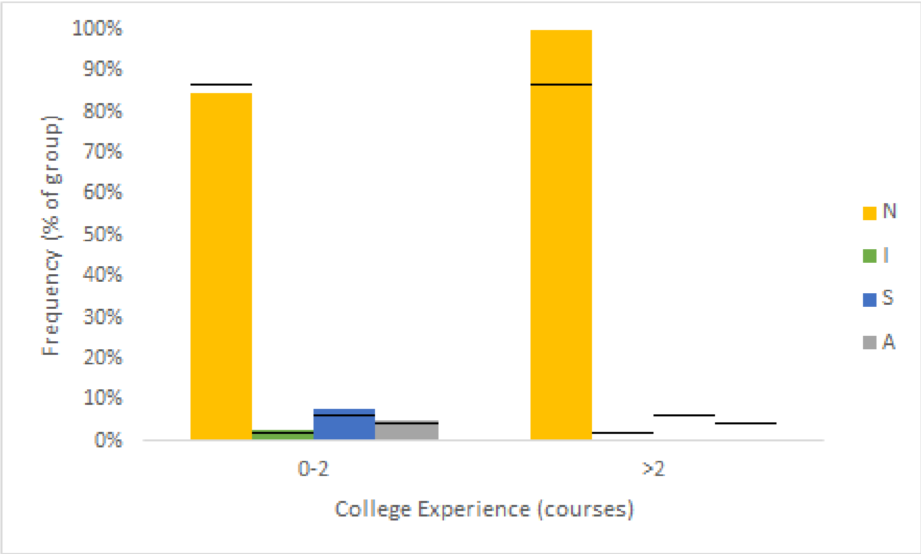

Figure 34. Degree of transfer of knowledge about vertical motion by college experience group on Question 6 of the pretest. Experience is based on student responses to Question 22 on the pretest. See the caption for Figure 12 for more information.

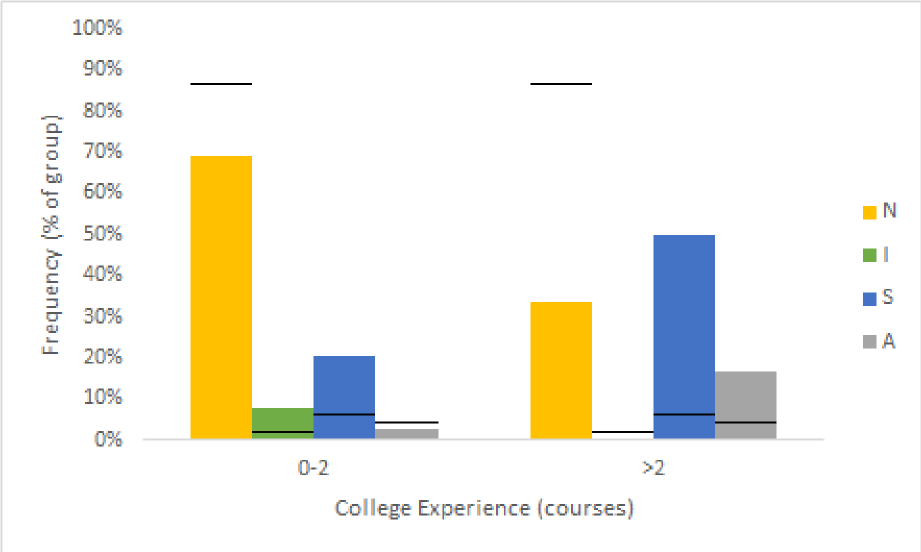

Figure 35. Degree of transfer of knowledge about vertical motion by college experience group on Question 7 of the Post-Exam I survey. Experience is based on student responses to Question 17 on the Post-Exam III survey. Expected values are based on pretest data. See the caption for Figure 12 for more information.

Results of Short-Answer Analysis - Quantitative

Because of the small sample size of the multi-method analysis, I did not conduct any inferential statistical tests on the short-answer response data.

For discussing the results of the descriptive quantitative portion of the short-answer analysis, I will first focus solely on the data that take into consideration a student's total physical science background experience (see Table 21). From those data, the mode code for the 0-2 experience group was invariably "N," which means that students with 2 or fewer course-years of previous physical science background most frequently did not engage in successful and accurate transfer of the measured concepts. The code "A" represents responses of which the degree of transfer was not clear, so it is very possible to argue that some "A" responses were really "N" responses or that some "A" responses were really "I" or "S" responses. To account for this, I used two different approaches to incorporating "A" responses into the results. The first, the "conservative approach," considered what would happen to the observed modes if all "A" responses were actually unidentified "N" responses; the second, the "liberal approach," considered what would happen if all "A" responses were actually unidentified "S" responses. The modes for the 0-2 experience group are never affected by taking the conservative approach nor by taking the liberal approach (see Figures 12-19). While the expected modes based on calculated expected frequencies from the pretest were indeed "N," the failure to engage in transfer still occurred more frequently than expected among the 0-2 experience group in all concept categories on the pretest. By contrast, the 0-2 experience group exhibited "N" responses more frequently than expected on the Post-Exam I survey when discussing T-ρ, but otherwise "N" responses on the Post-Exam I survey were less frequent than expected. "I" responses were invariably less frequent (or non-existent) on the pretest but always more frequent than expected on the Post-Exam I survey. "S" responses were always less frequent than expected, with the exception of responses involving ρ on the Post-Exam I survey; however, in every concept category there was still growth in "S" responses between the pretest and Post-Exam I survey. The concept category with the least variable degree of transfer on the pretest was ρ ($v_{pre,obs} \approx 0.0526$), meaning that students were the most uniform in this case in their inability to transfer ideas about density; all concept categories were otherwise equally variable ($v_{pre,obs} \approx 0.1053$). The concept category with the least variable degree of transfer on the Post-Exam I survey was V ($v_{post,obs} \approx 0.2105$) and the most variable was ρ ($v_{post,obs} \approx 0.5263$), which was ten times as variable on the Post-Exam I survey as it was on the pretest. The latter reveals a shift from "N" towards successful transfer (accurate or otherwise) of ideas about density, even if it did not ultimately change the mode. Responses were generally less variable than expected across all concept categories on the pretest and more variable than expected on the Post-Exam I survey, in comparison to the calculated expected variation ratios. The interesting exception to this rule were the responses to T-ρ on the Post-Exam I survey, which were more variable than expected.

The mode codes for the >2 experience group were no different from those for the 0-2 experience group on the pretest. However, on the Post-Exam I survey, the modes for the >2 experience group were "S" in all concept categories other than T-ρ. The widespread appearance of "S" as a mode on the Post-Exam I survey for the >2 experience group will be a pattern also exhibited in the secondary school and college experience analyses. Unlike students with 0-2 course-years of total physical science background experience, students with more than 2 course-years under their belt most frequently successfully engaged in accurate transfer of ideas about temperature, density, and vertical motion, deviating from those categories' expected modes of "N." The frequency of each degree of transfer was more or less the opposite of that of the 0-2 experience group: “N” was always less frequent than expected, “I” compared to the expected frequency differently depending on the test and concept category, and “S” was always more frequent than expected. If the conservative approach is taken of assuming all “A” codes are “N” codes, the mode for ρ on the Post-Exam I survey becomes “N”; otherwise, no modes are affected by the conservative approach. If we take the liberal approach of assuming all “A” codes are “S,” the modes for T-ρ on both the pretest and the Post-Exam I survey become an “S.” The lowest variation ratio on both tests was that of V ($v_{pre,obs} \approx 0.1429$, $v_{post,obs} \approx 0.4286$). While V was the least variable concept category on both tests, it was also notably the concept category with the greatest change in variability between tests. T-ρ was the most variable concept category on the pretest and ρ was the most variable on the Post-Exam I survey, both having a variation ratio of $v_{pre,obs} \approx 0.6429$. The >2 experience group was more variable than expected in their degrees of transfer across all concept categories and tests. The degrees of transfer exhibited by the >2 experience group were always more variable than those exhibited by the 0-2 experience group.

We now shift our attention to the quantitative impact of just secondary school experience on degree of transfer (see Table 22). Students with 2 years or fewer of past physical science experience in secondary school, like in the case of combined total experience, had a mode code of “N” for all concept categories in all tests, identical to the expected mode. Students consistently failed to engage in transfer on the pretest more than expected, with the interesting exception of V (see Figures 20-27), a phenomenon that will be elaborated on in the Discussion. In that case, the observed frequency of failure to engage in transfer (f_{obs} = 24 responses) was just barely less than the expected frequency ($f_{exp} \approx 24.7353$ responses). “N” on the Post-Exam I survey was always less frequent than expected. The observed frequency of “I” was either greater or less than expected depending on the test and concept category. The observed frequency of “S” was typically less than the expected frequency on the pretest and greater than the expected frequency on the Post-Exam I survey, but interestingly it was slightly more frequent than expected when looking at ρ (f_{obs} = 2 versus $f_{exp} \approx 1.7059$) and *V* (f_{obs} = 2 versus $f_{exp} \approx 1.7059$) on the pretest. Turning all "A" codes into "N" codes or "S" codes has no effect on any mode. The least-variable concept categories on the pretest were *T* and ρ ($v_{pre,obs} \approx 0.1379$). *V* was the least variable on the Post-Exam I survey ($v_{post,obs} \approx 0.2414$). The concept category on the Post-Exam I survey that was the most variable was ρ ($v_{post,obs} \approx 0.5172$), a rather sizeable change in variability compared to the pretest. Variation ratios were smaller than expected on three concept categories on the pretest, but *V* on the pretest ($v_{pre,obs} \approx 0.1724$ versus $v_{pre,exp} \approx 0.1471$) was more variable than expected, bringing to mind the odd quirks with how its observed frequencies of codes compared with its expected frequencies. All concept categories were more variable than expected on the Post-Exam I survey.

While I will report the modes and variation ratios for the >2 experience group, it is extremely important to highlight that results involving the >2 secondary school experience group should be taken with a grain of salt due to its small sample size, which only amounted to five students. These five students generally deviated from the expected modes. As expected, the modes for *T*, ρ, and *V* on the pretest were "N." However, *T*-ρ had a mode of "S," although this was a very variable concept category ($v_{pre,obs}$ = 0.6000) and indeed the most variable concept category on the pretest. On the Post-Exam I survey, *T*, *T*-ρ, and *V* had "S" for their modes, whereas ρ was bimodal with a mode of both "S" and "A." The five students failed to engage in transfer less frequently than expected in almost every case, but interestingly "N" was more frequent than expected for *V* on the pretest (f_{obs} = 5 versus $f_{exp} \approx 4.2647$). As before, there were an equal number of cases in which there were more "I" codes than expected and cases in which there were fewer “I” codes than expected. Successful and accurate transfer was generally more frequent than expected; however, on the pretest the code “S” was less frequent than expected for ρ and *V* ($f_{obs} = 0$ versus $f_{exp} \approx 0.2941$ for both concept categories). Taking the conservative approach and considering all “A” codes as failed transfer would make ρ on the Post-Exam I survey and *T*-ρ on the pretest bimodal with both “N” and “S” as the modes. Taking the liberal approach and considering all “A” codes as successful and accurate transfer would make ρ unimodal with a mode of “S” on the Post-Exam I survey but would otherwise produce no change in results. As previously mentioned, *T*-ρ was the most variable concept category on the pretest. The least variable category was *V*, in which all five students failed to exhibit any transfer ($v_{pre,obs} = 0$). *V* on the pretest was the only concept category that was less variable than expected ($v_{pre,obs} = 0$ versus $v_{pre,exp} \approx 0.1471$). On the Post-Exam I survey, *T* and *T*-ρ were tied for the most variable concept categories ($v_{post,obs} = 0.4000$) and ρ and *V* were tied for the least variable concept categories ($v_{post,obs} = 0.2000$). All concept categories on the Post-Exam I survey were more variable than expected. In general, the degrees of transfer exhibited by the >2 experience group were more variable than the 0-2 experience group on the pretest and less variable than the 0-2 experience group on the Post-Exam I survey. The exceptions to this rule are *V* on the pretest and *T*-ρ on the Post-Exam I survey.

When looking at the results when only considering college experience (see Table 23), please keep in mind that this sample was larger than the other two ($n = 45$) but there were still only six students in the >2 experience group. Bringing our attention to the 0-2 experience group first, “N” was always the mode regardless of concept category and test, as expected. The code “N” always occurred more frequently than expected on the pretest (see Figures 28-35), with the exception of *V* on the pretest ($f_{obs} = 33$ versus $f_{exp} = 33.8000$), and less frequently than expected on the Post-Exam I survey. Uniquely, the "I" code across both surveys was less frequent than expected for the *T* and ρ concept categories but more frequent than expected for the *T*-ρ and *V* concept categories. "S" always occurred on the pretest less frequently than expected except for *V* on the pretest (f_{obs} = 3 versus f_{exp} = 2.6000). It was always more frequent than expected on the Post-Exam I survey. The conservative approach of treating all "A" codes as "N" codes and the liberal approach of treating all "A" codes as "S" codes produce no change to any mode. *T* and *T*-ρ were tied for the most variable concept category on the pretest ($v_{pre,obs} \approx 0.2564$), whereas ρ and *V* were tied for the least variable ($v_{pre,obs} \approx 0.1538$). ρ was the most variable concept category on the Post-Exam I survey ($v_{post,obs} \approx 0.4872$) and *V* was the least ($v_{post,obs} \approx 0.3077$). The concept categories were always less variable than expected on the pretest, with the exception of *V* on the pretest ($v_{pre,obs} \approx 0.1538$ versus $v_{post,obs} \approx 0.1333$), and more variable than expected on the Post-Exam I survey.

On the pretest when looking at the >2 experience group, the mode for three of the concept categories was the expected "N," but ρ had a mode of "I." None of the modes for the concept categories on the Post-Exam I survey were as expected: *T*, *T*-ρ, and *V* had "S" for their modes, and ρ was trimodal with "N," "S," and "A" being the modes. The "N" code was less frequent than expected in all concept categories across both tests, apart from *V* on the pretest (f_{obs} = 6 versus f_{exp} = 5.2000). "I" was less frequent than expected in all concept categories across both tests except for *T* (f_{obs} = 1 versus $f_{exp} \approx 0.5333$) and ρ (f_{obs} = 3 versus $f_{exp} \approx 0.5333$) on the pretest. The "S" code was more frequent than expected in all concept categories across both tests except for *V* on the pretest (f_{obs} = 0 versus f_{exp} = 0.4000). The conservative approach of treating all "A" codes as "N" codes causes *T*, *T*-ρ, and *V* on the Post-Exam I survey to be bimodal with the modes "N" and "S," and also causes ρ on the Post-Exam I survey to be unimodal with the mode

"N." The liberal approach of treating all "A" codes as "S" codes causes T-ρ on the pretest to be bimodal with the modes "N" and "S," and also causes ρ on the Post-Exam I survey to be unimodal with the mode "S." The two most variable concept categories on the pretest were ρ and T-ρ ($v_{pre,obs} = 0.5000$) and the least variable concept category was V, which was the only concept category on the pretest that was less variable than expected ($v_{pre,obs} = 0.0000$ versus $v_{pre,exp} \approx 0.1333$). On the Post-Exam I survey, T, T-ρ, and V all had observed variation ratios of 0.5000, whereas in ρ the results were evenly distributed across all three of the concept category's modes. As a consequence, ρ was the only concept category on the Post-Exam I survey that was less variable than expected ($v_{post,obs} = 0.0000$ versus $v_{pre,exp} \approx 0.2222$). V on the pretest and ρ on the Post-Exam I survey were the only concept categories for the >2 experience group that were less variable than the 0-2 experience group.

Results of Short-Answer Analysis - Qualitative

I will next report the anecdotal observations I made from the qualitative analysis. The first, most general question I pursued during the qualitative analysis was whether there was a noticeable difference in responses as one moved from students with low experience to students with high experience on Appendix D.1. What I was able to conclude after a spot of pseudoscientific pattern searching was that differences do exist but that the degree to which those differences are pronounced depends on the type of background experience being measured. When considering the total experience of the students, two noticeable changes in *response quality* (by which I mean the frequency that accurate transfer was exhibited and the depth to which physics concepts were elaborated) occurred when moving from 0 course-years to more than 3 course-years. The first was a subtle improvement in quality between responses by students with 0 to 1.5 course-years of physical science experience and those by students with 2 course-years of physical science experience. Short sentences and an absence of any tangible answer pervaded both sub-groups but the responses by students with 2 course-years of physical science experience were generally a little more on the mark and thought-out than those by students with fewer than 2 course-years of experience. Here are three examples of pretest responses by the 0-2 total experience group that I consider representative of the group as a whole:

- **Student #107 (1 course-year):** *Heat causes plates to move and thus causes major shifts like Earthquakes. Not sure about density.*
- **Student #86 (2 course-years):** *no clue*
- **Student #51 (2 course-years):** *I think when a rock gets heated enough it travles away from the heat and probabley all that shifting of rocks has some affect on plate tectonics. as for density im not quite sure maybe denser rocks are harder to move or break.*

Much more dramatic was the improvement in quality between responses by students with 2 course-years of physical science experience and those by students with more than 2 course-years of experience. The latter group exhibited much more in-depth discussion of the quantitative relationships between properties. The following are three responses by the >2 total experience group that I feel exemplify the group:

- **Student #34 (2.5 course-years):** *Temperature differences create a convective flow due to materials having different densities at different temperatures.*
- **Student #38 (More than 3 course-years):** *Colder water is more dense than warmer water. When the water warms it becomes less dense and "pushes" against the Earth surface with a lesser amount of pressure. The decrease in pressure caused by the rise in*

temperature allows the Earth to shift it's plates. I believe this causes natural disasters like tsunamis.

- **Student #60 (More than 3 course-years):** *the temperature changes due to the proxmity to the magma that is in the core of the earth. The densities change depending on the material of that layer. Possibly if the layer is hotter it would be less dense because the particles are moving faster.*

Qualitative differences when just looking at secondary school experience or college experience were less clear. When only considering secondary school experience, there was a noticeable increase in response quality between the 0-2 and >2 experience groups on the Post-Exam I survey, but differences in responses to the pretest were not as clear. There seems to be an increase in quality between the two experience groups on the pretest, but I cannot confidently rule out the observer-expectancy effect, particularly concerning the 0-2 experience group. Likewise, there seems to be a slight increase in quality between the 0-2 and >2 experience groups on the pretest when just considering college experience, but this may again be due to my own observer-expectancy effect. There seems to be little difference between the two groups on the Post-Exam I survey, which is at odds with the quantitative portion of the analysis. I believe any heterogeneity on the Post-Exam I survey is likely in part due to the appearance of near transfer of material learned in the introductory earth science course.

The word *convection*, which I argued is strong evidence of transfer from an academic physical science setting, is explicitly invoked only by students with more than 2 course-years of total physical science experience, appearing in 2 (14%) of that experience group's pretest responses and 4 (28%) of that group's Post-Exam I survey responses. Searching for *convection* in responses when only considering secondary school or college experience, however, leads to a surprising puzzle which will be further toyed with in the Discussion section. *Convection* is explicitly invoked in 2 responses on the pretest and 2 responses on the Post-Exam I survey by students with 2 years or less of physical science experience in secondary school (6.9% of that experience group); it is not explicitly invoked at all in any response to the pretest by students with more than 2 years of physical science experience in secondary school, but is invoked in 2 responses on the Post-Exam I survey (40% of responses by that experience group). The term appears in 2 (5.1%) of responses to the pretest by students that reported taking 2 or fewer physics courses in college and in 6 (15%) of their responses to the Post-Exam I survey. By comparison, it does not show up at all in pretest responses by students who reported taking more than 2 physics courses in college and appears twice in their responses to the Post-Exam I survey (17% of that experience group).

Because only concept categories directly relevant to plate tectonics were measured in the quantitative portion of the analysis, it is interesting to look at instances in which transfer of other physics concepts occurred that were or were not relevant. The two most frequent of these concepts were *pressure* (not directly relevant to how temperature and density drive plate tectonics) and *thermal expansion* (which was an aspect of T-ρ that students did not necessarily have to touch on). The frequency of transfer of these two concepts (regardless of accuracy) are tabulated in Table 24. There are a few other specific cases however that are worth bringing up. Student #121 (who reported having 1 year of physical science secondary school experience and having previously taken 1 physics course in college on the pretest, for a total of 2 course-years of total experience) exhibited transfer of ideas about the relationship between mass and phases of matter (specifically boiling point) on the pretest. Student #18, in addition to bringing up thermal expansion, transferred ideas about a relationship between density and rate of motion on the pretest; they reported 2 years of secondary school experience and more than 3 course-years of total experience on the pretest. Student #43 did not explicitly bring up thermal expansion, but did transfer the idea that as an object's temperature increases the speed of its particles increases on the pretest; they reported having taken more than 3 physics courses in college and having more than 3 course-years of total experience on the pretest. Student #69 explicitly connected mass and density on the Post-Exam I survey; they reported 1 year of secondary school experience and having taken 1 physics course in college on the pretest. Student #54 exhibited transfer of ideas about energy on the Post-Exam I survey, specifically about the Sun as a source of heat energy for the Earth and about the Earth having its own internal energy; they reported 1.5 years of secondary school experience and that they had not previously taken any physics courses in college on the pretest. It should be noted that introductory earth science courses sometimes teach students about Earth's energy sources, so Student #54's apparent transfer of ideas about energy may likely be an instance of near transfer of earth science content rather than far transfer of physical science content. Lastly, Student #74 transferred ideas about the relationship between density and heat capacity on the pretest; they reported having previously taken 1 physics course in college on the pretest.

Frequency of Miscellaneous Physics Concepts on Question 6 of Pretest and Question 7 of Post-Exam I Survey According to Reported Previous Physical Science Experience

		CONVECTION			
		PRETEST		**POST-EXAM I SURVEY**	
Experience Type	Amount of Experience	*f* (# responses)	*p* (% of experience group)	*f* (# responses)	*p* (% of experience group)
Total Experience	0-2 course-years	0	0%	0	0%
	>2 course-years	2	14%	4	28%
Secondary School Experience	0-2 years	2	6.9%	2	6.9%
	>2 years	0	0%	2	40%
College Experience	0-2 courses	2	5.1%	6	15%
	>2 courses	0	0%	2	17%

		PRESSURE			
		PRETEST		**POST-EXAM I SURVEY**	
Experience Type	Amount of Experience	*f* (# responses)	*p* (% of experience group)	*f* (# responses)	*p* (% of experience group)
Total Experience	0-2 course-years	2	11%	1	5.2%
	>2 course-years	2	14%	1	7.1%
Secondary School Experience	0-2 years	4	14%	2	6.9%
	>2 years	0	0%	0	0%
College Experience	0-2 courses	2	5.1%	1	2.6%
	>2 courses	2	33%	1	17%

Table 24. *F* requency at which students with a particular amount of physical science background experience invoked certain concepts in their responses to Question 6 of the pretest and Question 7 of the Post-Exam I Survey. *f* represents the number of responses that appeared in which a particular concept was invoked. *p* is the percent of students within an experience group who invoked a particular concept in their responses. Keep in mind that the total number of analyzed responses varied between experience types.

Frequency of Miscellaneous Physics Concepts on Question 6 of Pretest and Question 7 of Post-Exam I Survey According to Reported Previous Physical Science Experience					
		THERMAL EXPANSION			
		PRETEST		POST-EXAM I SURVEY	
Experience Type	Amount of Experience	*f* (# responses)	*p* (% of experience group)	*f* (# responses)	*p* (% of experience group)
Total Experience	0-2 course-years	0	0%	0	0%
	>2 course-years	2	14%	0	0%
Secondary School Experience	0-2 years	1	3.4%	0	0%
	>2 years	1	20%	0	0%
College Experience	0-2 courses	2	5.1%	0	0%
	>2 courses	0	0%	0	0%

Table 24 (cont.)

During my qualitative analysis, I identified the student with ID number 112 (who was only included when considering just college experience) as a likely outlier. Despite reporting themselves as having previously taken no physics courses in college on the pretest, their responses to the pretest and Post-Exam I survey were nearly perfect and exhibited a lot of successful transfer, data that conform much more closely with the >2 experience group. I do not believe this to have significantly affected the results of the quantitative analysis.

One last point that I noted during my qualitative analysis that may be of interest to researchers pursuing other research questions was that there seemed to be a clear relationship between how much background experience students had and how confident they were about warranting a guess for the short-answer question. When looking at total experience, 6 or 32% of the 0-2 experience group reported having "no clue" or simply left their answer blank on the pretest and 5 or 26% of students (more than a quarter) answered such on the Post-Exam I survey. By comparison, there was only one instance on the pretest and one instance on the Post-Exam I survey in which a student in the >2 experience group did not attempt to answer the question (7.1% of that experience group). This discrepancy was much more dramatic when just looking at secondary school experience—*only* students in the 0-2 experience group were not confident enough to attempt an answer, with 5 or 17% of pretest responses and 6 or 21% of Post-Exam I survey responses being either "No clue" or left blank. When just looking at college experience, results were a little more mixed. Again, only students in the 0-2 experience group were not confident enough to attempt an answer on the pretest, with 7 or 18% of those students not even guessing, but both groups hosted students who were not confident enough to answer the question on the Post-Exam I survey (6 or 15% of students in the 0-2 experience group versus 1 or 17% of students in the >2 experience group).

While not seemingly relevant to this study, it may be worth mentioning for use in other studies or to satisfy the interest of instructors that the post-test responses in general to me were more focused and exhibited greater control of geoscientific language than the pretest responses. There does not appear to be any relationship between these characteristics and past physical science experience, however.

4. DISCUSSION

Complications

As previously mentioned, while a large number of students took the pre-test and post-test surveys, the actual sample sizes for my analyses were smaller than the total number of participants. Of the 114 students who took the pretest and 119 students who took the Post-Exam III survey, only 103 (≈90.35%) and 63 (≈79.75%), respectively, responded to the demographics questions and thus could be included in the analyses. Fifty-seven students who took the Post-Exam I and Post-Exam II surveys did not make it to the Post-Exam III survey to report their demographics information, and so their responses also could not be included in the analysis of the post-test multiple-choice data nor that of the short-answer data. Forty-two students were not included on the pretest multiple-choice data, pre-vs.-post multiple-choice data, nor short-answer data analyses because they did not take the pretest, although some of those students did not qualify for analysis in general. Furthermore, some students who did take the Post-Exam III survey did not answer all of the multiple-choice questions on the Post-Exam I survey, resulting in only 56 students (≈70.89% of the total 79 students who took the Post-Exam III survey) being analyzed for Post-Exam I survey Question 2 and only 53 students (≈67.09% of the total 79 students who took the Post-Exam III survey) being analyzed for Post-Exam I survey Question 4. Because I limited the short-answer analysis to students who reported the same past physical science experience on the pretest and on the Post-Exam III survey, only 14 students (≈17.72% of the total 79 students who took the Post-Exam III survey) were analyzed for Question 7 on the Post-Exam I survey. It is thus worth mentioning that it is unknown whether the excluded students' responses to the questions of interest could have affected the results had they reported their background experience.

After removing those students who did not qualify for analysis, the sample sizes for the multiple-choice and short-answer analyses were so small that it was difficult to make concrete conclusions about the data. The power of the chi-squared test increases with larger frequencies in each cell of the contingency table so it is possible that with a larger sample size, more null hypotheses from the multiple-choice analyses may have been able to be rejected. The frequency values in a contingency table can be maximized by minimizing the number of rows and columns in the table. The frequency values for the multiple-choice analyses were maximized by dividing students who answered the demographics questions into two bins and comparing the frequency of students in each experience group who answered a particular question correctly with that of those who answered a particular question incorrectly, producing a 2×2 contingency table. While this did increase the sizes of each cell's frequency values, they were still too small to be greatly confident about the conclusions of the chi-squared tests. It would be very interesting to see a similar study be performed with a larger sample size so that the original idea for a 3×4 contingency table could be implemented.

Another notable complication for the study was the wording of the questions. As the survey was not designed with my transfer study in mind, certain questions' wordings potentially biased the student responses to exhibit what appears to be far transfer. The chief offender was Question 6 on the pre-test and the equivalent Question 7 on the Post-Exam I survey. Although the wording for the two questions is slightly different, both explicitly ask how temperature and density relate to plate tectonics. For some students, the words *temperature* and *density* may have acted as hints to think about physics. As outlined in Bassok & Holyoak (1989) and Detterman (1993), many studies have found that transfer only very rarely occurs in *H. sapiens* in the absence of some explicit hint or suggestion to use one's prior knowledge (e.g. Reed, Ernst, & Banerji, 1974). It is thus possible that some students may have been biased to talk about temperature and density because those words were present in the question. It is also possible that the wording may have led some students to deceivingly invoke the words *temperature* and *density* without actually engaging in transfer. While it was difficult to weed these students out given that I could not ask them about their responses, I decided that at the very least simply using the words *temperature* and *density* in a response was not enough for it to qualify as an "S" response in those two categories (often it was also not enough to qualify for an "I"). Questions 2 and 4 on the pre-test and Post-Exam I survey were sort of on the other end of the spectrum, in which no sort of hint was provided to encourage students to think about the problems from a physical science perspective. It is possible that the absence of any hints biased students towards not engaging in successful and accurate transfer that they might have otherwise engaged in.

While not a complication per se, it is possible that some differences between the results of the multiple-choice analysis and the short-answer analysis may be due to students having to explain their answers in the short-answer question. As mentioned in the Introduction, explaining one's reasoning has been found to be a good way to force transfer to happen (Brown & Kane, 1988).

As this study focused on the effects of reported *physical science* experience on understanding of novel earth science concepts, students' previous background experience in *earth science* was not taken into account in my analyses. Student earth science experience was reported in Questions 20 and 21 on the pretest and Questions 15 and 16 on the Post-Exam III survey. Because this was an introductory earth science course, most students had little previous earth science experience, but there was at least one student who reported having taken more than

3 college earth science courses on the Post-Exam III survey. This study on its own cannot rule out earth science experience as a confounding variable.

Findings

The first question to ask before all others is, “Did transfer even happen?” Some such as Detterman would be skeptical, as mentioned in the Introduction. I think the answer is a firm “Yes,” and I can back that up by pointing to the chi-squared tests of independence that I ran on the multiple-choice data. With the current data, there are three possibilities that we can rule out with confidence. First, we can statistically rule out the independence of college experience and accuracy on Question 2 of the pretest (see Table 6). This means that whether a student was correct or incorrect on Question 2 of the pretest depended on the number of physical science courses previously taken by that student in college. Second, we can statistically rule out the independence of secondary school experience and accuracy on Question 2 of the Post-Exam I survey (see Table 10). This means that whether a student was correct or incorrect on Question 2 of the pretest depended on the years that student spent learning about physical science in secondary school. Third, we can also statistically rule out the independence of total experience and accuracy on Question 2 of the Post-Exam I survey (see Table 8). It is very interesting that the chi-squared test of independence was not able to rule out the independence of college experience and accuracy on Question 2 of the Post-Exam I survey (see Table 12); I will return to this discrepancy shortly. The other lines of evidence for transfer in this population comes from the short-answer data, which display blatant examples of transfer (see Appendix D.1). In particular, the prevalence of ideas like *convection*, *pressure*, and *thermal expansion* among others in responses to Question 6 of the pretest makes for very compelling evidence that transfer occurred.

Few chi-squared tests were able to rule out the possibility that students with little to no previous physical science experience engaged in just as much or just as little transfer as students with a lot of previous physical science experience. We must thus then posit the question, "Why were so few null hypotheses rejected?" The first possibility is that it is an artifact of the small sample size. With a larger sample size, it is possible that more chi-squared tests would indicate their null hypotheses should be rejected. In addition to low sample size, though, there's another factor that may be contributing to the failure to reject so many null hypotheses, and it is through the few cases where the null hypotheses could be ruled out that we can infer what may have happened. Looking at Tables 6, 8, and 10, we can see that a student's previous physical science experience in college affected whether they answered Question 2 of the pretest correctly, and that a student's total previous physical science experience and just their previous physical science experience in secondary school affected whether they answered Question 2 of the Post-Exam I survey correctly. However, even though the observed frequencies deviated statistically significantly from what would be expected if both experience groups engaged in the same amount of transfer, no null hypothesis could be ruled out for any chi-squared test for proportionality. Curious, is it not, that one test suggests there absolutely is a correlation between experience and accuracy but the other cannot decipher any sort of difference? I believe this discrepancy signals that other factors masked any transfer of physical science knowledge, and that such a mask covers all the multiple-choice data analyses, only being thin enough for chi-squared tests of independence to conclusively identify a correlation when looking at past college experience on the pretest and past total and secondary school experience on the Post-Exam I survey. I thus conclude that the amount of background experience one has in a relevant subject is but one of many influencing variables that affected a student's accuracy when responding to the multiple-choice questions, rather than being the sole determining variable for the students' accuracy. It is possible that studies that fail to provide evidence of transfer in *Homo sapiens* do so in part because they fail to control for these other influencing variables.

We can also see this on Table 21. On the pretest, students most frequently failed to engage in transfer regardless of their total physical science experience. The variation ratios tell a richer story, however[3]. If we imagine what a population's code frequency would look like as more and more of their responses exhibited successful and accurate transfer, at first the entire population would fail to exhibit one instance of transfer. Then, slowly, more and more instances of transfer would be exhibited by the population so that, while "No transfer" was still the mode, the code frequencies of the population would be a lot more heterogenous than before. Little by little, as successful and accurate transfer becomes more and more frequent, the population would still be heterogenous but "successful and accurate transfer" would now be the mode. Finally, the population would ultimately be homogenous again with all responses exhibiting successful and accurate transfer. This exact sort of transition can be seen when comparing the responses of students with little to no total physical science experience on Question 6 of the pretest to the responses of students with a lot of total physical science experience: while the modes are the same between the two groups, the high experience group is a lot more heterogenous in its code frequencies. It is possible that past experience is responsible for this increase in heterogeneity.

The directionality of Figures 2 and 4 seem to support this conclusion that past experience plays a small role—however, it must be emphasized that this is conjecture as in most cases the differences between the performance of each experience group was not statistically significant.

[3] The variation ratios on Tables 22 and 23, while also interesting, are too heavily impacted by the small sample size of the high experience group to contribute meaningfully to this discussion.

Interestingly, the lack of directionality and the generally smaller differences between each experience group's performance on Figures 3 and 5 suggest that Question 2 may have possessed helpful cues to stimulate transfer or otherwise guide students to the right answer that Question 4 lacked. This would conform with observations that *H. sapiens* often require some sort of hint to encourage transfer (Detterman, 1993). Question 2 was about the motion of the mantle at a subduction zone, where the subducting slab stuck out underneath the overlying plate and into the asthenosphere, having a general appearance of moving downwards. Question 4 was about the motion of the mantle at a divergent boundary, which may have been harder for students to conceptualize. The more-difficult spatial concepts involved in Question 4 may not have facilitated transfer, which might be why not one null hypothesis involving Question 4 could be ruled out by chi-squared tests.

Not all evidence of greater transfer among the high experience group was vague, however. Figures 12, 14, 16, and 18 perhaps show the clearest evidence that an increasing amount of past experience facilitates transfer. There is a very visible difference between the degrees of transfer exhibited by students with little past experience and those exhibited by students with a lot of past experience on the pretest. While the mode degree of transfer—no transfer—is the same for both experience groups, the high experience group was more successful in engaging in transfer of ideas about temperature, density, and the relationship between these two properties than the low experience group, as evidenced by the higher percent frequencies of the “I” and “S” codes in the former. This is also generally true for the more-anecdotal Figures 20, 22, 24, 26, 28, 30, 32, and 34.

One possibility that was posed during the course of the data analysis was that the effects of having past physical science experience may not manifest themselves until after starting the earth science course. In other words, regardless of whether evidence of far transfer shows up on the pretest, having more physical science background may make it easier to learn about geophysics and thus evidence of far transfer may show up on the post-test. My research has no means of weeding far transfer from near transfer on the post-test questions, so I cannot declare my findings conclusive, but there are certainly pieces of evidence that support and do not support the hypothesis. Quantitatively, there was a large difference in responses by low experience groups to the short-answer question compared to responses by high experience groups on the Post-Exam I survey (see Figures 12-35). While both experience groups see a reduction in "N" responses and an increase in "I" and/or "S" responses on the Post-Exam I survey, this reduction is far more dramatic in the high experience group, especially when just looking at secondary school experience or college experience. In those anecdotal cases, "N" responses often completely disappear in the high experience group on the Post-Exam I survey compared to the same group on the pretest. However, it should absolutely be noted that the small size of the high experience groups in those two types of experience certainly has a hand in exaggerating the differences between columns. If we look at the more-reliable analyses based on total experience where the sample sizes of the two experience groups are more similar, though, we can still see evidence that having more previous physical science experience promoted transfer by the time students took the Post-Exam I survey. There was a striking increase in the frequency that students with a lot of past total experience engaged in successful and accurate transfer of ideas about temperature and vertical motion, both compared to the low experience group on the Post-Exam I survey and compared to themselves on the pretest. More evidence in support of the idea that experience promotes transfer is on display in Tables 21-23, where we can see that the low experience groups invariably had a mode of "N" in every concept category on both the pretest and the Post-Exam I survey, meaning that most responses failed to exhibit any transfer, but that the high experience groups almost always shifted from a mode of “N” on the pretest to a mode of “S” on the Post-Exam I survey.

Not all of my evidence supports the conclusion that past experience manifests itself on post-tests. The chi-squared tests performed on the pre vs. post analysis (see Tables 14-19), for example, generally do not agree with the hypothesis. For the most part, I cannot statistically rule out the possibility that there was no difference in how students with a lot of physical science experience changed their answers between the pretest and the Post-Exam I survey compared to students with little to no physical science experience. Likewise, while quantitatively there was clear improvement in the high experience groups between Question 6 of the pretest and Question 7 of the Post-Exam I survey (see Tables 21-23), qualitatively this did not seem to be the case when looking specifically at college experience.

Surprisingly, when looking at just college experience on Table 18, the results of the chi-squared test of proportionality contradict my above statement and suggest that there *is* a statistically significant difference between the frequency at which students in the low experience group answered Question 2 incorrectly on the pretest but correctly on the Post-Exam I survey and the frequency at which students in the high experience group did so. This is the only test of proportionality in my entire study to reject a null hypothesis, and it is even more perplexing because the results of the corresponding test of independence was unable to rule out that there was any difference between the low and high experience groups in how they changed their responses. While the results of this test of proportionality on its own support the idea that past experience would lead students with low experience to perform differently from students with high experience, I actually caution the reader from gathering any conclusions based on this finding. Similar to the analysis of the short-answer responses, when looking just at college experience the size of *n* for the high experience group is very low (in this case, only 9 students out of 74). I think this makes Table 18 anecdotal at best, and it is entirely plausible that with a higher sample size the results of the chi-squared tests on Table 18 would be very different. It is interesting to note, however, how low the *p* value for the test of independence was and how close it came to indicating that the null hypothesis for the test of independence should be ruled out. I would be very interested to see how a larger sample size would affect this, and whether the low sample size of the high experience group in particular is causing a Type II error with the test of independence or a Type I error with the test of proportionality.

Another morsel of interest is the discrepancy between the results of the chi-squared tests for independence on the pretest multiple-choice data (see Tables 2-7) and those on the post-test multiple-choice data (see Tables 8-13). The effects of far transfer are more evident when looking at total experience and just secondary school experience on the Post-Exam I survey (see Tables 8 and 10) than they are on the pretest (see Tables 2 and 4), as we can rule out the independence between accuracy and past physical science experience on the Post-Exam I survey. It may be that students engaged in transfer on both tests, but it is interesting that the chi-squared tests can only pin down the presence of far transfer on the Post-Exam I survey. It is the complete opposite when looking at past college experience (see Table 6); the null hypothesis could only be ruled out on the pretest, not the Post-Exam I survey (see Table 12). My interpretation of these discrepancies is that secondary school physical science experience promoted learning through positive far transfer, whereas college physical science experience hindered learning through negative far transfer; between the two types of background experience, the effects of secondary school experience overpower the effects of college experience when looking at the effects of total experience. Why this is is a mystery to me. It harkens back to my discussion in the Introduction about how students may have more difficulty transferring secondary school experience compared to college experience. This could be food for thought to inspire further research on transfer in relation to experience.

In going over my results of the short-answer question, it is difficult to not be struck by the strange exceptions on Figures 26 and 34 surrounding transfer of ideas about vertical motion. In describing how I qualitatively analyzed student responses to Question 6 on the pretest and Question 7 on the Post-Exam I survey, I discussed how usage of the specific word *convection* is a good indicator of transfer of concepts from physical science courses. One would thus predict that use of the word *convection* correlates with past physical science experience. This was very evident when considering a student's total past physical science experience, but it was not at all obvious when considering just secondary school or just college experience. In fact, *convection* showed up in the language of students with very low secondary school experience on the pretest but not at all in the language of students with a lot of secondary school experience, and results were similar when looking at college experience. I would argue though that this does not challenge the idea that *convection* is an indicator of physical science experience. The fact that *convection* is completely restricted to the high experience group when looking at total experience to me indicates that *convection* appeared in the low experience group when looking at secondary school experience or when looking at college experience because those students who used the word convection still had a lot of total experience overall. Because use of the word *convection* qualified a response for a code of "S" (successful and accurate transfer) in the vertical motion concept category, the abundance of *convection* in the low experience group when looking at secondary school or college experience is likely what for example caused the observed frequency of “N” (no transfer) to be lower than expected and the observed frequency of “S” to be higher than expected on Figure 34.

Another term that appeared where it was least expected in responses to Question 6 on the pretest and Question 7 on the Post-Exam I survey was *pressure*. Pressure is not a driver of plate tectonics, so one would expect students with a lot of physical science experience to not transfer ideas about pressure. However, it was more often than not the case that high experience groups exhibited transfer of ideas about pressure more frequently than low experience groups did (see Table 24). This finding supports the small but growing amount of research that suggests that some misconceptions about pressure are picked up in traditional physics courses (Kuethe, 1991). Interestingly, while the percentage of students in the high total experience group who brought up pressure was higher than that of students with little to no total past experience, it is worth noting that discussions of pressure among the high experience group were richer and more quantitatively-minded than the simplistic mentions of pressure by the low experience group, suggesting that transfer of pressure while misguided was indeed still facilitated by having more experience. This correlation between past experience and response quality was not reflected when just looking at college experience, which similarly had the high experience group discuss pressure more frequently than the low experience group. I cannot identify where these misconceptions about pressure originate. However, perhaps secondary school experience holds the key. Uniquely among the types of experience, pressure was only ever brought up by students with little to no secondary school experience in physical science. This may suggest that secondary school has a unique opportunity to weed out misconceptions about pressure that college courses do not have.

Thermal expansion is an important detail of the relationship between temperature and density, but the short-answer questions did not necessarily require it in order to specify how temperature and density drive plate tectonics—"warm things are less dense than cold things" was satisfactory. Because ideas about thermal expansion can be presumed to require some physical science experience, I thought it worth noting which student responses actually specifically laid out how temperature relates to volume. Explicit ideas about thermal expansion only came up in two responses, so any conclusions that can be made from the data are weak at best, but ideas about thermal expansion were indeed only reported by students in the high total experience group. Such a pattern was not present when just looking at secondary school experience or college experience.

As far as the interesting instances of specific students who exhibited transfer of irrelevant physical science concepts other than pressure are concerned, there does not appear to be any notable pattern in their amounts of experience.

Implications

What may have contributed to the results I observed, and how might educators in the future use my research to inform their craft? I would like to take some time here to consider the results in the broader picture of transfer and the application of transfer research to pedagogy.

A similar study that is worth bringing up is one published in 2000 by Sadler and Tai. In their research, they compared the performance in an introductory college physics course of undergraduate students across the U.S. who did take Physics in high school to the performance of those who did not take Physics in high school. They found that the mean grade in the college physics course for students who did take Physics in high school was higher than the mean grade for students who did not take Physics in high school (Sadler & Tai, 2001). This can be taken to support the idea that transfer of secondary school experience to a college setting occurs within a single discipline and that having more secondary school experience facilitated learning over the course of a class. My results suggest that something similar is true for transfer between disciplines but that such transfer is more difficult for students to accomplish, conforming to Barnett & Ceci (2002)'s taxonomy.

Perhaps the way in which students received their background experience caused them to respond in a particular fashion. In seeing that the quantity of background experience is but one influencing variable in how students will respond to cross-discipline questions—and a very small one at that—perhaps it may be better for education higher ups to emphasize *quality* of teaching over *quantity* when preparing students for a particular college discipline. Emphasizing quality may help increase the depth at which students learn material, which facilitates transfer in human children (Karbach & Kray, 2009) and adults (Schliemann & Magalhães, 1990).

Beyond just quantity vs. quality, is there a style of teaching that encourages far transfer? Fuchs et al. (2003) designed a third-grade curriculum that focuses on abstraction and metacognition in order to assess whether far transfer could be facilitated by explicitly teaching students to apply their knowledge (in this case of problem-solving skills in math) to novel scenarios. The ability for students to construct broad schemata was addressed by having students sort novel math problems by what made them similar and what made them different—for example, two problems may be sorted together if they were identical in how they should be solved but differed in key vocabulary. Students were explicitly taught what *transfer* was and that when faced with a novel problem they should try to sort it into one of the bins that they had learned to sort problems with similar solutions into. The authors found that when presented with unfamiliar tests containing unfamiliar questions proctored by unfamiliar people, students who had received the curriculum focused on abstraction and metacognition performed significantly better and engaged in successful and accurate far transfer much more frequently than those students whose otherwise-identical curriculum did not focus on abstraction and metacognition (Fuchs, et al., 2003). Barnett and Ceci (2002) summarize other studies that have shown that encouraging metacognitive processes in class increases the likelihood of successful far transfer. Doing so may be particularly important in courses such as Physics in which students struggle more to apply what they learn to other things on their own (Bassok & Holyoak, 1989).

5. CONCLUSION

The results of this study in general seem to suggest that past experience plays some minor role in far transfer but that it is only one of many variables that influence whether a human student's brain will take material learned in one setting and apply it to a new setting. There are some clear examples of differences between students with little experience and students with a lot of experience, but the two groups are not as distinguishable by their responses as common sense would predict.

While the earth sciences build off of the physical sciences, taking physics courses before taking an earth science course does not necessarily predict transfer of that physical science knowledge to geophysics. When transfer does occur, college experience seems to be more useful with novel questions while secondary school experience seems to be more useful over the long run. This means that educators should focus on better encouraging transfer in secondary school physical science courses, such as by highlighting examples of ways that that information can be transferred to other disciplines, as suggested by (Fuchs, et al., 2003). It may also be beneficial for earth science educators to highlight places where experience with other disciplines may be helpful, in accordance to research on the mechanisms facilitating transfer such as (Gick & Holyoak, 1983).

The degree of ecological validity of this study is greatly complicated by the small sample size of the data, and so while I am proud for my results to be an early stepping stone in the study of transfer and background experience, I encourage the results to be taken with a grain of salt and for other researchers to further pursue this subject.

REFERENCES

Barnett, S. M., & Ceci, S. J. (2002, July). When and where do we apply what we learn?: A taxonomy for far transfer. *Psychological Bulletin, 128*(4), 612-637. doi:10.1037/0033-2909.128.4.612

Bassok, M., & Holyoak, K. J. (1989). Interdomain transfer between isomorphic topics in algebra and physics. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*(1), 153-166. doi:10.1037/0278-7393.15.1.153

Brown, A. L., & Kane, M. J. (1988, October). Preschool children can learn to transfer: Learning to learn and learning from example. *Cognitive Psychology, 20*(4), 493-523. doi:10.1016/0010-0285(88)90014-X

Brown, A. L., Kane, M. J., & Long, C. (1989, October/December). Analogical transfer in young children: Analogies as tools for communication and exposition. *Applied Cognitive Psychology, 3*(4), 275-293. Retrieved July 27, 2017, from http://onlinelibrary.wiley.com/journal/10.1002/(ISSN)1099-0720

Cooper, G., & Sweller, J. (1987, December). Effects of schema acquisition and rule automation on mathematical problem solving transfer. *Journal of Educational Psychology, 79*(4), 347-362. doi:10.1037/0022-0663.79.4.347

Detterman, D. K. (1993). The case for the prosecution: Transfer as an epiphenomenon. In D. K. Detterman, & R. J. Sternberg (Eds.), *Transfer on trial: Intelligence, cognition, and instruction.* Norwood, New Jersey, United States of America: Ablex Publishing Corporation. Retrieved October 8, 2017

Fuchs, L. S., Fuchs, D., Prentice, K., Burch, M., Hamlett, C. L., Own, R., . . . Jancek, D. (2003, June). Explicitly teaching for transfer: Effects on third-grade students' mathematical problem solving. *Journal of Educational Psychology, 95*(2), 293-305. doi:10.1037/0022-0663.95.2.293

Gick, M. L., & Holyoak, K. J. (1983, January). Schema induction and analogical transfer. *Cognitive Psychology, 15*(1), 1-38. doi:10.1016/0010-0285(83)90002-6

Karbach, J., & Kray, J. (2009, November). How useful is executive control training?: Age differences in near and far transfer of task-switching training. *Developmental Science, 12*(6), 978-990. doi:10.1111/j.1467-7687.2009.00846.x

Kendall, M. G. (1938, June). A new measure of rank correlation. *Biometrika, 30*(1-2), 81-93. doi:10.1093/biomet/30.1-2.81

Kotovsky, K., Hayes, J. R., & Simon, H. A. (1985, April). Why are some problems hard? Evidence from Tower of Hanoi. *Cognitive Psychology, 17*(2), 248-294. doi:10.1016/0010-0285(85)90009-X

Kuethe, D. O. (1991, January). Confusion about pressure. *Physics Teacher, 29*(1), 20-22. Retrieved April 2, 2018

Kurt, S., & Ehret, G. (2010). Auditory discrimination learning and knowledge transfer in mice depends on task difficulty. *Proceedings of the National Academy of Sciences, 107*(18), 8481-8485. doi:10.1073/pnas.0912357107

Pearson, K. (1900, July). 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. *Philosophical Magazine Series, 50*(302), 157-175. doi:10.1080/14786440009463897

Reed, S. K., Ernst, G. W., & Banerji, R. (1974). The role of analogy in transfer between similar problem states. *Cognitive Psychology, 6*, 436-450. doi:10.1016/0010-0285(74)90020-6

Rosenthal, R., & Fode, K. L. (1963, April 1). Psychology of the scientist: V. Three experiments on experimenter bias. *Psychological Reports, 12*(2), 491-511. doi:10.2466/pr0.1963.12.2.491

Royer, J. M. (1979). Theories of the transfer of learning. *Educational Psychologist, 14*(1), 53-69. doi:10.1080/00461527909529207

Sadler, P. M., & Tai, R. H. (2001). Success in introductory college physics: The role of high school preparation. *Science Education, 85*(2), 111-136. Retrieved August 29, 2017, from http://users.nber.org/~sewp/events/2004.05.28/SadlerAndTai_RoleOfHighSchoolPhysics.pdf

Salomon, G., & Perkins, D. N. (1989). Rocky roads to transfer: Rethinking mechanism of a neglected phenomenon. *Educational Psychologist, 24*(2), 113-142. doi:10.1207/s15326985ep2402_1

Sautter, V., Toplis, M. J., Wiens, R. C., Cousin, A., Fabre, C., Gasnault, O., . . . Wray, J. J. (2015, July 13). In situ evidence for continental crust on early Mars. *Nature, 8*, 605-609. doi:10.1038/NGEO2474

Schliemann, A. D., & Magalhães, V. P. (1990). Proportional reasoning: From shopping to kitchens, laboratories, and, hopefully, schools. In G. Booker, P. Cobb, & T. N. de Mendicuti (Ed.), *Proceedings of the Annual Conference of the International Group for the Psychology of Mathematics Education with the North American Chapter 12th PME-NA Conference. 3*, pp. 67-73. Oaxtepec, México: International Group for the Psychology of Mathematics Education. Retrieved January 27, 2017, from http://www.pmena.org/

BIOGRAPHY OF THE AUTHOR

Graham Hummel-Hall is a graduate student enrolled in the Master of Science in Teaching program at the University of Maine. He was born in Greenwich, Connecticut, but spent most of his life living in Arizona and Wisconsin. He briefly attended the University of Wisconsin-Stevens Point pursuing a bachelor's degree in biology, but transferred after a year to the University of Wisconsin-Madison. Graham graduated in May 2015 with a Bachelor of Science degree after majoring in zoology, focusing specifically on vertebrate paleontology. While most of his classes focused on biology and geoscience, he also took many classes on the French language and linguistics. Throughout his time in Madison, Graham worked at the University of Wisconsin-Madison Geology Museum as a tour guide and introduced people of all ages to the wonders of Earth's biotic and abiotic history. He also did volunteer work in the fossil prep lab on reassembling *Triceratops horridus* fossils and interned in the museum's repository. Currently, Graham substitute teaches at secondary schools in Bangor and in Old Town, Maine. After completing his studies at the University of Maine, he plans to teach Biology in high school. When not studying or working, Graham spends most of his time drawing, composing music, playing video games, and making his wife smile. He is a passionate animal lover, a devoted fan of animation, and frequently enjoys contemplating the tree of life and how humans fit into the universe. He is a candidate for the Master of Science in Teaching degree from the University of Maine in May 2018.