2019

Exploring the relationship among students' preconceptions, attitudes, and major

Matthew Dwyer

ABSTRACT

EXPLORING THE RELATIONSHIP AMONG STUDENTS' PRECONCEPTIONS, ATTITUDES, AND MAJOR

Matthew Dwyer, M.S.
Department of Physics
Northern Illinois University, 2019
Michael Eads, Director

The Force Concept Inventory (a multiple-choice classical mechanics exam given as a pre-test within the first two weeks of a course and as a post-test within the last two weeks of a course) is one of the most widely used tools in physics education research. The most basic research, which has been implemented at Northern Illinois University for many years, employs a pre- and post-test structure to observe improvements in students' scores over a semester and is used as a quantitative representation of students' learning over the duration of a course. While many studies have been published about best practices to improve results on the Force Concept Inventory between the pre-test and post-test, fewer studies have been conducted on students' thinking upon entering a course. Often, students' initial thoughts on classical mechanics are full of preconceptions. Some of these preconceptions are intuitive ways of observing the world, but they don't often work the way students might imagine.

In this research, the relationship among students' preconceptions upon entering a course, their declared majors, and whether or not they view physics in the way an expert in the field would (as defined by the Colorado Learning Attitudes about Science Survey) is explored. This relationship is used to discover which common preconceptions are held by different

groups of students and could be used to steer the various courses to teach towards correcting those preconceptions specifically. This method of teaching towards correcting preconceptions has been shown to be effective in various research studies, but there has been little work done to see if preconceptions differ by student. This research aims to fill that gap in knowledge.

"Expert-like" students who enrolled in calculus-based classical mechanics courses designed for physics and engineering majors did not have unique preconceptions aside from a misapplication of the term "net" force. Non-"expert-like" students who enrolled in algebra-based classical mechanics courses designed for other science majors, as well as algebra-based courses in general physics for all other majors, did have preconceptions unique to their group, mainly that "if motion requires force, then forces must generate motion." Both groups share very common preconceptions, most notably the preconception that an object in motion must have a force acting upon it in the direction of motion.

NORTHERN ILLINOIS UNIVERSITY
DE KALB, ILLINOIS

MAY 2019

EXPLORING THE RELATIONSHIP AMONG STUDENTS' PRECONCEPTIONS, ATTITUDES, AND MAJOR

BY

MATTHEW DWYER

A THESIS SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE

MASTER OF SCIENCE

DEPARTMENT OF PHYSICS

Thesis Director:
Michael Eads

ACKNOWLEDGEMENTS

I'd like to thank Dr. Michael Eads, who never once balked at my pursuing a wholly original idea without any sort of guarantee I'd find any conclusions worth looking into. Hopefully my conclusions have, in fact, been worth looking into.

As well, thanks to Dr. Roland Winkler, who helpfully wrote the "niuthesis" package in LaTeX I've used to write this thesis, making it look more professional than I've ever been.

Patrick Coyle helped with the statistical analysis and was patient when I asked the same question multiple times in a row without realizing it.

Dr. Cynthia Campbell instructed me in performing education research and graded a first draft of this thesis with much higher marks than I probably deserved.

Additional thanks to Elaine and Tom Dwyer for always being supportive and only being marginally disappointed the nights I bailed on dinner to work more on my thesis. I promise I'll be around for more meals now that I've finished this.

Hannah Dwyer - Thanks for being a great older sister and only beating me up when I *really* deserved it. I hope you're as proud of me as I am of you.

Tim Brandt and Lucas Hendren - Two of my best friends and best support system. Thank you for never being mad when I last-minute canceled our Dungeons and Dragons sessions and for telling me when my ideas were trash. You're right - they were. I couldn't ask for better friends.

Cody Sauber, thanks for always encouraging me when I was stressing about my work and letting me complain to you about my insignificant problems and treating them like they were actually important. The afternoons we spent sitting around your place watching old clips of Bo Burnham were the pinnacle of my summer. I'm thrilled to have more time available for those. Drinks are on me this month.

Ralph Helm - For being a fantastic roommate who understood that I put off doing the dishes to de-stress and never once complaining to me about it. You took one for the team. Thank you for your patience and warmth.

Jacob Pearson - For always taking the time to hit me up when you're in Sycamore. I promise I'll be able to finally go out and gallivant with you now. You've been a good friend for a decade, and I'm so happy to have you around.

Lastly - Brianna Stamas - Thanks for being the best part of my day and buying me coffee, two equally important things. While finishing this degree is exhilarating, I'd still say our crossing paths has been the highlight of my time at NIU. Have fun at CERN.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Preconceptions

One issue with teaching classical mechanics that makes it unique to other fields of physics is that the concepts discussed make up the bulk of how the world is experienced. Students will oftentimes come into a class with some general intuition about how objects *should* behave based on prior experiences with kinematics, momentum, or forces, although the lived examples were never given such formal names. These experiences will be referred to as "throwing a ball," "taking a turn on a bike a bit too sharply," or "getting jostled by an older sibling." It is these intuitive preconceptions that can often lead to incorrect solutions within course problems.

An example of a very prevalent preconception is the notion that a falling body's speed is proportional to its weight [1, 2]. This preconception doesn't have any bearing in rigorous study but instead is formulated from the student's interactions with the outside world. One such physical example I had observed as a child was dropping a brick and then dropping a feather - clearly mistaking air resistance for the apparent difference in the force of gravity acting upon the bodies.[1] This line of thinking has occasionally been referred to as "Aristotelian physics"[2]. To quote Heron's take on such preconceptions or "difficulties":

[1] To be a bit pedantic here, it should be noted that I was right. A brick will, in fact, fall faster than a feather in any everyday circumstance. While this is an imperfect example, it does illustrate how hard it can be to blend Newtonian mechanics with lived experiences. The issue doesn't always arise from what's observed but with the combination of observables and theory.

"[A difficulty is] the use of a specific idea or pattern of reasoning instead of those we consider correct and appropriate. The notion that "closer means stronger" might be a value-neutral resource. The assumption that the brightness of lightbulbs in a circuit depends on their proximity to the battery is a difficulty."[3]

In the above situation, the original observation of "a brick falls faster than a feather" could naturally lead to the incorrect explanation presented. These preconceptions often arise when a student attempts to explain new content presented in a formal course by way of ideas or observations that are already familiar [1, 2].

Another common preconception is the idea that motion *requires* a force [1]. This can be observed by a pushed object slowing down and requiring further prodding to continue along its path. Of course, the reduction in speed is caused by frictional forces, but to a novice, friction is often overlooked because of a lack of knowledge of its existence or a lack of understanding of how friction actually works. It is here that a student would have confusion reconciling $\sum_i F_i = ma$ with this observation and can lead to further confusion on what a force or acceleration actually is. The same preconception can be applied to many different situations, all with incorrect results.

Once a student has these preconceptions, it is incredibly hard to change a student's mental model [4]. [2] Once preconceptions are addressed, it can pave the way to a more interconnected way of thinking about physics, instead of having students view it as just "classroom knowledge" [4–6].

[2]Part of the reason we develop these theories is based in human evolution and our innate ability to notice patterns for survival. While wholly unrelated to this study, I spent quite a bit of time diving into psychology papers on this topic because I was so fascinated with it.

1.2 "Knowing Just Enough to Be Dangerous"

While the idea of these preconceptions is already fairly well known, what hasn't been studied nearly as extensively is if these preconceptions can differ by student, instead of a "writ large" attitude, in particular, the case when students with some general knowledge or interest of physics will enter a classical mechanics course with different preconceptions than their peers.[3] For example, students with no knowledge of friction can't make an error in applying it to a problem since it isn't in their physicist's toolbox, whereas peers with a weak grasp on the concept can apply it to disastrous results. This general idea of "knowing just enough to be dangerous" isn't wholly original - it's closely related to the famed Dunning-Kruger effect, defined as:

> *"...a cognitive bias in which people of low ability have illusory superiority and mistakenly assess their cognitive ability as greater than it is. The cognitive bias of illusory superiority comes from the inability of low-ability people to recognize their lack of ability. Without the self-awareness of metacognition, low-ability people cannot objectively evaluate their competence or incompetence."* [7]

However, the "enough to be dangerous" concept differs in that, instead of students with less knowledge being more confident in their response, students with more (rough) knowledge simply open themselves up to more elaborate errors in thinking. As a final anecdotal example of this topic, one can go to some resources about the "Flat Earth conspiracy" and bear witness to the jumps from physics principles to ridiculous conclusions. One such grouping of connections is the usage of reference frames and Newton's second law. That acceleration

[3]While preconceptions exist across all fields, classical mechanics was chosen for study because it's how every person experiences the world, giving the most opportunity for a student to come up with explanations for phenomena.

we feel because of gravity's pull is *actually* the Earth accelerating upwards - just a switch of reference frame and $F = ma$.[4]

The fact that this misguided way of thinking is entering cultural discourse will, ideally and unfortunately, make this concept of "knowing just enough to be dangerous" easier to understand.

1.3 Purpose Statement

The purpose of this study is to explore the relationship among students' preconceptions about classical mechanics upon entering a course, their attitudes regarding physics as a whole, and whether they're enrolled in a class geared towards physics majors and engineers or a more general education course.

1.4 Research Questions

The research questions for this study are:

- Do students enrolled in introductory physics classes intended for physical science and engineering majors view physics more as an expert in the field would when compared to peers enrolled in a classical mechanics course meant for students of other disciplines?
- Do students with a more "expert-like" view on physics have different preconceptions than students with less "expert-like" views upon entering the course, shown through different common erroneous responses on the Force Concept Inventory (FCI) [8] ?

[4] This is an actual concept of some importance in general relativity, which only bolsters the point of a *little* knowledge being a dangerous thing.

1.5 Hypothesis

Students enrolled in a classical mechanics course intended for physics and engineering majors *will* view physics more like an expert in the field would; otherwise, they wouldn't have an initial interest in the class. Because many of these courses are required for given majors (physical sciences, engineering), the student has already shown to be in a major that involves these topics. Assuming this, more "expert-like" students could also have some prior knowledge about higher level topics, but probably not enough detail to apply those topics correctly. This inability to utilize theories, but the knowledge to attempt to use them, will cause students enrolled in classical mechanics courses intended for physics and engineering majors to have different incorrect responses to questions.

1.6 Key Concepts

Attitudes and Beliefs: Feelings regarding how well one may do within science courses, or how a student feels physics impacts one's life and the importance of the subject as a whole [9].

"Expert-like" Understanding of Physics: Viewing physics the way that someone who practices it everyday might [9], for example, viewing physics as a series of interconnected theories and phenomena instead of simple formulas that need to be memorized to pass a course.

Preconceptions: Certain alternative ways of reaching conclusions, usually based on observed phenomena [4]. An example outside the field of physics is the theory of spontaneous generation [10]. One example of spontaneous generation is the observation that a person leaving out spoiled meat will eventually find maggots have started feasting on it, causing the person to believe that the maggots sprang up from the meat itself. The act of maintaining these preconceptions is sometimes referred to as “belief perseverance” [4].

Most Common Distractor: When giving a multiple-choice test to all students, the value being looked at is the most common wrong answer for a given question included to *distract* the students from the correct response with competing, inaccurate theories. Thus the term “distractor” [11].

CHAPTER 2

METHODS

2.1 Participants

Participants in this study were university students enrolled in physics courses, including both algebra- and calculus-based courses. All participants in this study were volunteers, and participation was not mandatory or linked to any sort of grade for the course, although participation was encouraged. Variables such as gender, age, race, or income were not studied. The relevant information included courses the student was enrolled in (calculus-based courses are generally geared towards engineers and physics majors, whereas algebra-based courses tend to be for pre-professional students) and the student's major, if declared. Figures for breakdowns by major between PHYS 253 and PHYS 210 + PHYS 150 are in Appendix B.

While it would be ideal if every student within a course took the FCI and the Colorado Learning Attitudes about Science Survey (CLASS), it is not an accurate assumption that all students will participate. All participants attended Northern Illinois University between Fall 2015 and Spring 2017 and enrolled in one of the following physics courses: 150, 210, or 253. These classes are defined in Section 3.1. While both the FCI and CLASS are given as a pre- and post-test, the majority of analysis was done on the pre-test, before significant instruction has occurred. The post-test was briefly analyzed for discussion on how students' preconceptions change over the course of a semester, but will mostly be left for future work.

Students were also provided with an informed consent form wthat explains that their results may be used for course improvement and research purposes. Only students who consented have contributed data towards this study.

2.2 Materials

2.2.1 Colorado Learning Attitudes about Science Survey

The Colorado Learning Attitudes about Science Survey [9] is a 42-question Likert scale[1] survey designed to compare how a student feels about physics and science in general to how "experts" feel about the subject, where an expert is defined as a person who uses physics in their professional life, such as a professor or lab technician. An example question is: *"I think about the physics I experience in everyday life."* Clearly, an expert in the field would relay common everyday experiences to their studies, and if a student answers positively, they share that belief with an expert. At the end of the survey, a student's overall beliefs are compared to an expert's. The CLASS is graded by taking a student's percentage of expert-like responses and averaging over every student enrolled within a course to get an overall course expert-like percentage. The CLASS grading sheet also contains breakdowns by category of attitudes, which will not be explored within this study.

The CLASS has gone through multiple revisions to ensure questions make sense to students and their results are valid. The process for ensuring validity involved direct interviews with experts after taking the survey, interviews with students to confirm the questions made

[1] A Likert scale is defined as: *"The Likert Scale is a 5- or 7-point scale that offers a range of answer options — from one extreme attitude to another, like "extremely likely" to "not at all likely." Typically, they include a moderate or neutral midpoint"*[12]. All those surveys with, "How does X make you feel?" are on a Likert scale.

sense and the students' interpretations of what was asked of them were accurate, and a "factor analysis"[2] to breakdown statements into different categories. While most of the questions asked were interpreted correctly by students, ones that were not were either removed or reworked for the most recent version of the CLASS [13].

2.2.2 Force Concept Inventory

The FCI is the second instrument used within this study. The FCI is the most common concept inventory used in physics and has been utilized to provide a massive amount of data on physics education. It is a pre- and post-test given to introductory classical mechanics courses within the first two weeks of a class and the final two weeks. The FCI is a multiple-choice physics exam that asks questions to root out common errors beginning students might make or fundamental errors in conceptualizing physics. The FCI has been used generally to test teaching effectiveness,[3] but in this case it was used to determine a student's preconceptions upon entering a course by observing which questions students get incorrect and how. The questions have been designed to purposefully contain "distractors" - common incorrect ways of thinking - as possible solutions to test students' understanding of concepts.

2.3 Procedure

The FCI and CLASS were administered through differing means. The CLASS was an online test that was emailed to students and accessed through the Leaning About STEM

[2]While this type of correlational data analysis isn't utilized in my own study, there is a comprehensive source provided by Adams et al. [13].

[3]Many of these studies pertain to teaching students via "active methods," which usually means in a more "hands-on" approach as opposed to attending a lecture or teaching specifically to deconstruct students' preconceptions [14, 15].

Student Outcomes online tool via the Learning Assistance Alliance [16], whereas the FCI was taken during a physics lecture's adjoined laboratory course. In-person testing does have an increase in response rate over online testing, but those effects on response rate have not been studied within this research. Online tests generally have low response rates. The CLASS was administered to all introductory physics courses in the fall of 2018, but the FCI has data ranging back to 2013. The CLASS pre-test is within the first month of a course, the FCI pre-test is within the first week, and both CLASS and FCI post-tests are within the final two weeks of a course. Testing the first research question utilized the CLASS:

Do students enrolled in introductory physics classes intended for physics and engineering majors view physics more as an expert in the field would when compared to peers enrolled in a classical mechanics course meant for students of other disciplines?

For Fall 2018, the CLASS was given to the following courses at NIU: PHYS 150, PHYS 162, PHYS 180, PHYS 210, PHYS 211, PHYS 253, and PHYS 273, for a total student count of 124. Not every student who took the CLASS has their data used within this study. Students must answer a control question correctly (to ensure students are not selecting random responses) and must consent to having their responses be used for research. There was a number of students who took the CLASS but did not agree to share their data. Lastly, because FCI data was not analyzed for classes other than PHYS 253, PHYS 210, and PHYS 150, only those three courses contributed to the analysis within this study, bringing the total number of analyzed responses to 61.

Testing the second research question will utilized the FCI:

Do students with a more "expert-like" view on physics make different errors on introductory classical mechanics topics than their peers?

Using the data gathered from answering the first research question, students within PHYS 253, PHYS 210, and PHYS 150 were assumed to be experts or non-experts as a class. Because of an inability to connect individual students' replies on the CLASS to the FCI because of privacy concerns, it was assumed that every student within PHYS 253, PHYS 210, or PHYS 150 think the same way based on the majority within the course. A more individualized approach is suggested for future work for someone who has access to this unfiltered, non-anonymous data (not a graduate student).

Once the classes themselves were determined to be more expert-like or non-expert-like, all responses on the FCI were graded, and the most chosen incorrect answer for each question was noted. If a question had a different most chosen distractor between the two groups of students (expert vs. non-expert), it was flagged for further review and an attempt was made to explain the difference in responses. The logic that led to each incorrect response was discussed, although no interviews were done during this process. Instead, common preconceptions that have been explored in past research were analyzed.

2.4 Analysis

The official way of grading the CLASS is to employ the use of a rather unwieldy Excel file [17] to carefully look through every student's response, find the percentage of a student's responses that are considered consistent with experts in the field, then average every student within a course to calculate the overall course expert-like percentage. However, I found writing Python code to do this for me was much easier and enjoyable. Seeing as how the individual categories were not being analyzed, the overall percentage was easy enough to calculate by looking at the questions as being all equally weighted.

The FCI is more straightforward in its grading - each question has one correct response and four distractors. All the chosen distractors for a question were summed and listed with any question that had a different most commonly chosen distractor between experts and non-experts.

2.5 Study Limitations

As stated previously, it is not possible to link FCI and CLASS data back to a single student, as all data are received as anonymous responses. This has led to a few assumptions on my part, namely that a class as a whole will be composed of experts or non-experts. This is clearly incorrect, as every course will have a large breadth of students enrolled within it, but this assumption will hopefully serve as a jumping off point until this analysis can be run on unfiltered data. There is a potential way around this limitation, which will be left for future work. Instead of defining "expert-like" students by the CLASS, examine the ratio of total FCI questions correct to most chosen distractor for every student and question. This makes the assumption that "experts" always do better on the FCI, but it is worth pursuing.

CHAPTER 3

RESULTS

3.1 CLASS Responses

To test the first research question, "Do students enrolled in introductory classical mechanics classes intended for physics and engineering majors view physics more as an expert in the field would when compared to their peers enrolled in a general education course," the first step is to define which physics courses are intended for which majors. According to the NIU undergraduate course catalog, PHYS 150, PHYS 211, and PHYS 253 are described as the following[18]:

- **PHYS 150:** "Development of concepts and principles from selected topics in mechanics, electricity, heat, sound, and light. Application to everyday life. Not recommended for students who have had a year of high school physics. Not available for credit to students with credit in PHYS 150A."
- **PHYS 210:** "First semester of a two-semester sequence covering mechanics, heat, and sound. Includes lecture and laboratory sessions. Not available for credit to students with credit in PHYS 253."
- **PHYS 253:** "Physical laws governing motion, force, energy, rotation, and vibration using calculus. Primarily for majors in the physical and mathematical sciences and engineering. One three-hour laboratory a week. Not available for credit to students with credit in PHYS 210, PHYS 250, or PHYS 250A."

PHYS 253 is targeted towards physics and engineering majors, whereas PHYS 150 and PHYS 210 serve most other students. Also of note is that PHYS 210 is required for biology and health science majors, whereas PHYS 150 tends towards non-science majors. This could be a point worth examining more deeply in the future. The next step is to determine the percentage of "expert-like" behavior among students enrolled in the courses. This was done using the CLASS and graded for a single semester (Spring 2018). The results are shown in Table 3.1.

Table 3.1: Average Percentage Favorable by Class

Classes	Average Percentage Favorable	Participants
PHYS 150	46%	13
PHYS 210	53%	18
PHYS 253	60%	30
PHYS 150 and PHYS 210	50%	31

All data for the CLASS can be found in Appendix B. These values are surprisingly close together, with our naive guess being true - students enrolled in the courses set for physics and engineering majors do share more expert views than students in non-physical science majors. A Student's t-test was run on the above favorable percentages data to test for independence and had a p-value of 0.045, showing statistical significance of the difference between the two groups (PHYS 150 + PHYS 210 and PHYS 253). Because of an inability to match students to both their FCI and CLASS responses, I am unable to limit my FCI score comparisons between students who share expert views and students who do not. Instead of this precise exercise, FCI scores between the two groups of classes will have to suffice. However, it can be readily seen that while the number of experts in the courses geared for general education is regularly less than their engineer and physicist counterparts, there are still many non-experts in the more specialized courses. Because of this, some of the questions chosen for analysis on the FCI were in a bit of a middle ground between the two groups to allow for this lack of clear definition.

3.2 Numerical Results for All FCI Responses by Year and Semester

FCI pre-test data detailing students' responses to every question, as opposed to just a "number of questions correct" value, is available for the classes and semesters detailed in Table 3.2. The data itself can be found in Appendix B.

Table 3.2: Number of Student Responses to the FCI by Semester and Year

	PHYS 253 Pre-Test	PHYS 150 Pre-Test	PHYS 210 Pre-Test
Spring 2015	136	0	0
Fall 2015	214	0	137
Spring 2016	119	23	118
Fall 2016	195	50	131
Spring 2017	78	25	116

3.3 Determining Which FCI Questions Are Relevant

It is necessary to determine which questions differ in responses based on their grouping (experts and non-experts) rather than normal fluctuations. Here, "experts" refer to students within PHYS 253, and "non-experts" refer to students within both PHYS 210 and PHYS 150. Two separate tests were run in order to determine if the distractors chosen by group are dependent upon the group or random.

3.3.1 Questions Whose Total Responses Are Statistically Significant

First, all incorrect answers for a given question are assembled into a table (Table 3.3) for Question 2.

Table 3.3: Table Used for χ^2 Test for Independence

Q2	Distractor 1	Distractor 2	Distractor 3	Distractor 4
Experts	193	67	226	44
Non-Experts	155	42	246	45

Then, a χ^2 test for independence is run on each question. The questions that return a p-value of less than 0.05 are selected for further review[1] [19]. This test determines the likelihood that distractor responses are distributed differently between the two groups outside of random chance. In other words, the total responses are dependent upon the group we're observing.

3.3.2 Questions Whose Most Common Distractor Is Statistically Significant

The first test narrowed the FCI responses to questions that have statistically significant differences in all distractor responses. The second analysis will determine if the most common distractor is significantly different. This will answer the question, "Is the most common distractor for experts chosen with the same *frequency* for non-experts?" If, for example, both groups have the same most commonly chosen distractor, but experts have chosen it

[1] A detailed explanation of the χ^2 test is included in Appendix E.

90% of the time while non-experts chose it at 51%, that can still reveal a difference in thinking between the two groups.

The χ^2 test is run against the most common distractor for experts but should return the same findings if run against the most common distractor for non-experts, as it looks for significant changes in frequencies of responses. An example is Table 3.4, again for Question 2.

Table 3.4: χ^2 Test for Independence on Most Common Distractor

Q2	Number of Students Who Selected the Most Common Distractor for Experts	Number of Students Who Selected All Other Distractors
Experts	226	304
Non-Experts	246	242

3.4 Questions of Statistical Significance

Among all students across every year and semester, there are twelve questions that have statistically significant differences in distractor responses ($p < 0.05$), as shown in Appendix A.

Of these twelve questions, nine show a statistically significant difference in frequency of the most common distractor ($p < 0.05$; Figures 3.1 - 3.10). When going through each question in detail, instead of raw values, all tables will show percentages of students who selected that response, as it makes all tables more comprehensible. The correct FCI response is in bold in every table (Tables 3.5-3.13).

3.5 Uncertainty

When extrapolating outward from the subset of Northern Illinois University students who participated in this study to a more general population to estimate trends, there is some uncertainty present. To ensure all tables can fit and remain readable within the thesis, not every row was given an uncertainty. Instead, the maximum uncertainty has been calculated for a single situation and will be noted as an upper bound to show the validity of results.

Using $\sigma = \sqrt{\frac{p(1-p)}{n}}$ to calculate uncertainty in binomial statistics, p is the percentage of students who selected a certain response on the FCI and n is the total number of participants. With values of 50% and 550 students, this returns an error of 2.1%. The most any given FCI response percentage will fluctuate is 2.1%, in the least ideal case; 50% is chosen because it maximizes the numerator in the σ calculation.

3.5.1 Question 2

1. Two metal balls are the same size but one weighs twice as much as the other. The balls are dropped from the roof of a single story building at the same instant of time. The time it takes the balls to reach the ground below will be:
   (A) about half as long for the heavier ball as for the lighter one.
   (B) about half as long for the lighter ball as for the heavier one.
   (C) about the same for both balls.
   (D) considerably less for the heavier ball, but not necessarily half as long.
   (E) considerably less for the lighter ball, but not necessarily half as long.

2. The two metal balls of the previous problem roll off a horizontal table with the same speed. In this situation:
   (A) both balls hit the floor at approximately the same horizontal distance from the base of the table.
   (B) the heavier ball hits the floor at about half the horizontal distance from the base of the table than does the lighter ball.
   (C) the lighter ball hits the floor at about half the horizontal distance from the base of the table than does the heavier ball.
   (D) the heavier ball hits the floor considerably closer to the base of the table than the lighter ball, but not necessarily at half the horizontal distance.
   (E) the lighter ball hits the floor considerably closer to the base of the table than the heavier ball, but not necessarily at half the horizontal distance.

Figure 3.1: Questions 1 and 2 on the FCI

Table 3.5: Breakdown of Question 2 Responses by Student Subset

	Total Responses	Most Common Distractor	**A**	B	C	D	E
Experts	734	D	**28%**	26%	9%	31%	6%
Non-Experts	599	D	**19%**	26%	7%	41%	8%

p-value: 0.0156

Preconceptions of both groups:

Students of each group had a majority response of D. This distractor implies that an object's falling speed is proportional to its weight while also allowing for some uncertainty.

Preconceptions of experts:

Experts had a somewhat more even spread between B and D. Both distractors have the same implication (an object's falling velocity is proportional to its weight), but B doesn't have the same lack of specificity.

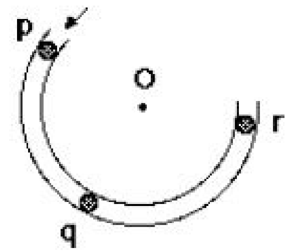
Preconceptions of non-experts:

Non-experts had their responses peak at D, with B a distant second (roughly 15% less). Because of the difference in wording between B and D, non-experts seem less confident in their reasoning, opting for the less certain of the two wordings.

3.5.2 Question 5

USE THE STATEMENT AND FIGURE BELOW TO ANSWER THE NEXT TWO QUESTIONS (5 and 6).

The accompanying figure shows a frictionless channel in the shape of a segment of a circle with center at "O". The channel has been anchored to a frictionless horizontal table top. You are looking down at the table. Forces exerted by the air are negligible. A ball is shot at high speed into the channel at "p" and exits at "r."



5. Consider the following distinct forces:
    1. A downward force of gravity.
    2. A force exerted by the channel pointing from q to O.
    3. A force in the direction of motion.
    4. A force pointing from O to q.

    Which of the above forces is (are) acting on the ball when it is within the frictionless channel at position "q"?

    (A) 1 only.
    (B) 1 and 2.
    (C) 1 and 3.
    (D) 1, 2, and 3.
    (E) 1, 3, and 4.

Figure 3.2: Question 5 on the FCI

Table 3.6: Breakdown of Question 5 Responses by Student Subset

	Total Responses	Most Common Distractor	A	**B**	C	D	E
Experts	736	C	3.53%	**16%**	32%	26%	22%
Non-Experts	599	C	9%	**11%**	45%	16%	20%

p-value: 1.51e-4

Preconceptions of both groups:

Students of each group had a majority response of C. This distractor implies that a force is required for an object to be in motion, but simultaneously ignoring the effects of support forces.

Preconceptions of experts:

Experts shared the same preconception as their peers, with the next most chosen distractor also including the support force effect.

Preconceptions of non-experts:

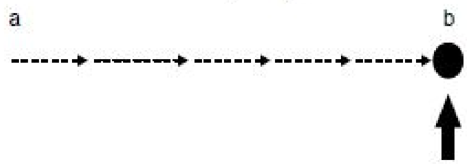
Non-experts had their responses dramatically peak at C, seemingly completely ignoring the fact that the channel exerts a force on the ball. Once again, supporting forces aren't recognized.

3.5.3 Question 11

Question 11 tangentially relies on information in Question 8. For clarity's sake, Question 8 has been included alongside Question 11.

USE THE STATEMENT AND FIGURE BELOW TO ANSWER THE NEXT FOUR QUESTIONS (8 through 11).

The figure depicts a hockey puck sliding with constant speed v_o in a straight line from point "a" to point "b" on a frictionless horizontal surface. Forces exerted by the air are negligible. You are looking down on the puck. When the puck reaches point "b," it receives a swift horizontal kick in the direction of the heavy print arrow. Had the puck been at rest at point "b," then the kick would have set the puck in horizontal motion with a speed v_k in the direction of the kick.



8. Which of the paths below would the puck most closely follow after receiving the kick?

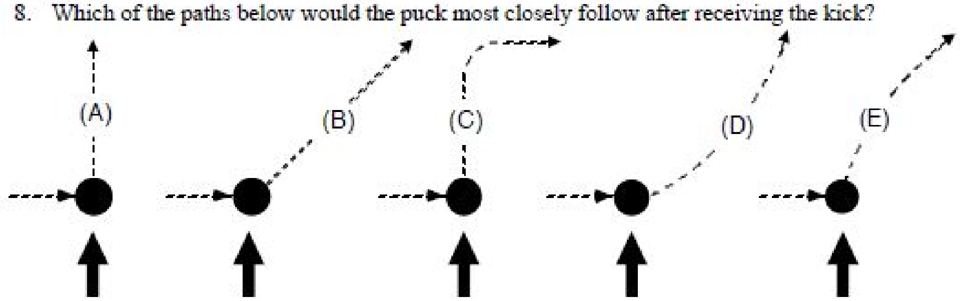


Figure 3.3: Question 8 and Figure on the FCI

11. Along the frictionless path you have chosen in question 8, the main force(s) acting on the puck after receiving the kick is (are):
    (A) a downward force of gravity.
    (B) a downward force of gravity, and a horizontal force in the direction of motion.
    (C) a downward force of gravity, an upward force exerted by the surface, and a horizontal force in the direction of motion.
    (D) a downward force of gravity and an upward force exerted by the surface.
    (E) none. (No forces act on the puck.)

Figure 3.4: Question 11 on the FCI

Table 3.7: Breakdown of Question 11 Responses by Student Subset

	Total Responses	Most Common Distractor	A	B	C	**D**	E
Experts	738	C	11%	27%	41%	**17%**	3%
Non-Experts	599	B	8%	41%	38%	**11%**	2%

p-value: 0.0159

Preconceptions of both groups:

Students of each group had a close to majority response of C. This distractor implies that a force is required for an object to be in motion but does include a support force.

Preconceptions of experts:

Experts' responses peaked sharply at C, with B roughly 13% behind. Both responses include the preconception that motion requires force.

Preconceptions of non-experts:

Non-experts were split nearly evenly between C and D. This, again, implies that non-experts routinely don't notice support forces.

3.5.4 Question 15

A large truck breaks down out on the road and receives a push back into town by a small compact car as shown in the figure below.

15. While the car, still pushing the truck, is speeding up to get up to cruising speed:
    (A) the amount of force with which the car pushes on the truck is equal to that with which the truck pushes back on the car.
    (B) the amount of force with which the car pushes on the truck is smaller than that with which the truck pushes back on the car.
    (C) the amount of force with which the car pushes on the truck is greater than that with which the truck pushes back on the car.
    (D) the car's engine is running so the car pushes against the truck, but the truck's engine is not running so the truck cannot push back against the car. The truck is pushed forward simply because it is in the way of the car.
    (E) neither the car nor the truck exert any force on the other. The truck is pushed forward simply because it is in the way of the car.

Figure 3.5: Question 15 on the FCI

Table 3.8: Breakdown of Question 15 Responses by Student Subset

	Total Responses	Most Common Distractor	**A**	B	C	D	E
Experts	735	C	**31%**	10%	53%	6%	1%
Non-Experts	600	C	**23%**	18%	49%	9%	1%

p-value: 2.8e-5

Preconceptions of both groups:

Both groups heavily favored C in their responses. This seems to be a variant of motion requiring force, wherein whatever object is generating the force will have a larger net force than anything interacting with that object.

Preconceptions of experts:

Experts' responses peaked sharply at C, the common preconception for both groups.

Preconceptions of non-experts:

Non-experts had nearly twice as many students respond with B than experts. This preconception implies that an object's resistance to motion is proportional to its weight, or some sort of misapplication of friction.

3.5.5 Question 21

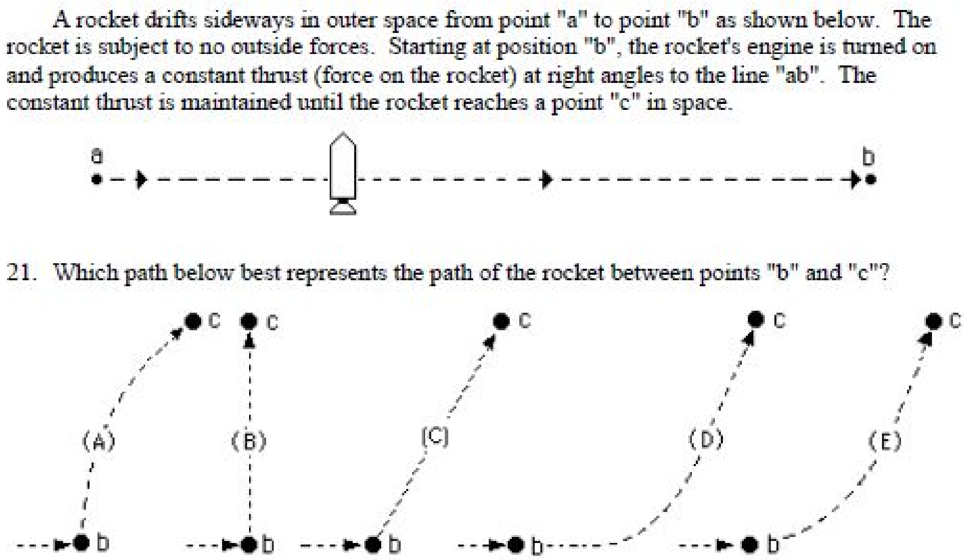


Figure 3.6: Question 21 on the FCI

Table 3.9: Breakdown of Question 21 Responses by Student Subset

	Total Responses	Most Common Distractor	A	B	C	D	**E**
Experts	735	C	10%	14%	24%	14%	**38%**
Non-Experts	597	B	11%	26%	24%	15%	**24%**

p-value: 0.04

Preconceptions of both groups:

Students of each group chose a distractor that implies an object experiencing a force has no buildup in a change of velocity; the object's speed is instantly changed, instead of a gradual rate from acceleration.

Preconceptions of experts:

Experts' most common distractor was C, implying a lack of gradual change of motion, but a sharp direction change and nothing else.

Preconceptions of non-experts:

Non-experts were split nearly evenly between B and C. The large amount of B responses seem to indicate a thought process that an object's past motion doesn't have any bearing on future motion; only the most recent force matters.

3.5.6 Question 25

25. A woman exerts a constant horizontal force on a large box. As a result, the box moves across a horizontal floor at a constant speed "v_o".

The constant horizontal force applied by the woman:

(A) has the same magnitude as the weight of the box.
(B) is greater than the weight of the box.
(C) has the same magnitude as the total force which resists the motion of the box.
(D) is greater than the total force which resists the motion of the box.
(E) is greater than either the weight of the box or the total force which resists its motion.

Figure 3.7: Question 25 on the FCI

Table 3.10: Breakdown of Question 25 Responses by Student Subset

	Total Responses	Most Common Distractor	A	B	**C**	D	E
Experts	734	D	3%	14%	**20%**	48%	15%
Non-Experts	597	D	5%	22%	**15%**	37%	20%

p-value: 6.45e-8

Preconceptions of both groups:

Students of each group chose D in a majority of responses. This follows the preconception line of thinking that motion requires a force in the direction of motion.

Preconceptions of experts:

Experts' responses peaked sharply at D, with the remaining percentage split somewhat evenly between the other responses (except for A).

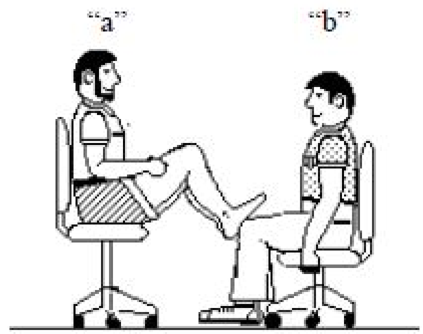
Preconceptions of non-experts:

Non-experts have a bit more even distribution between all responses, with two distractors mistaking weight for frictional force. Seemingly, non-experts conflate the two concepts.

3.5.7 Question 28

28. In the figure at right, student "a" has a mass of 95 kg and student "b" has a mass of 77 kg. They sit in identical office chairs facing each other.

Student "a" places his bare feet on the knees of student "b", as shown. Student "a" then suddenly pushes outward with his feet, causing both chairs to move.



During the push and while the students are still touching each other:

(A) neither student exerts a force on the other.
(B) student "a" exerts a force on student "b", but "b" does not exert any force on "a".
(C) each student exerts a force on the other, but "b" exerts the larger force.
(D) each student exerts a force on the other, but "a" exerts the larger force.
(E) each student exerts the same amount of force on the other.

Figure 3.8: Question 28 on the FCI

Table 3.11: Breakdown of Question 28 Responses by Student Subset

	Total Responses	Most Common Distractor	A	B	C	D	**E**
Experts	727	D	3%	10%	7%	48%	**32%**
Non-Experts	587	D	2%	20%	11%	53%	**14%**

p-value: 2.95e-3

Preconceptions of both groups:

Students of each group chose D a majority of the time. This is the second question where both groups tend to get confused with one object acting on another, where they tend to assume that the object that's the source of a force must output a larger force than it feels.

Preconceptions of experts:

Experts' responses were either D, which both groups chose, or the correct response.

Preconceptions of non-experts:

Non-experts also chose B with larger percentages than other responses. Distractor B applies the same preconception as D, but to a larger extent. (Person A not feeling any force at all, rather than feeling a smaller force.)

3.5.8 Question 29

29. An empty office chair is at rest on a floor. Consider the following forces:
    1. A downward force of gravity.
    2. An upward force exerted by the floor.
    3. A net downward force exerted by the air.

    Which of the forces is (are) acting on the office chair?

    (A) 1 only.
    (B) 1 and 2.
    (C) 2 and 3.
    (D) 1, 2, and 3.
    (E) none of the forces. (Since the chair is at rest there are no forces acting upon it.)

Figure 3.9: Question 29 on the FCI

Table 3.12: Breakdown of Question 29 Responses by Student Subset

	Total Responses	Most Common Distractor	A	**B**	C	D	E
Experts	729	D	18%	**44%**	2%	32%	4%
Non-Experts	586	A	32%	**32%**	5%	25%	7%

p-value: 6.058e-10

Preconceptions of both groups:

There was no majority preconception chosen by either group for Question 29.

Preconceptions of experts:

Aside from the correct response, experts chose D at much higher rates than any other response. This implies confusion regarding the term "net" force, a concept that someone with experience in physics might be familiar with in passing but unable to apply correctly.

Preconceptions of non-experts:

Non-experts, as has come to be expected, chose the distractor that ignores any sort of support force.

3.5.9 Question 30

30. Despite a very strong wind, a tennis player manages to hit a tennis ball with her racquet so that the ball passes over the net and lands in her opponent's court.

Consider the following forces:

1. A downward force of gravity.
2. A force by the "hit".
3. A force exerted by the air.

Which of the above forces is (are) acting on the tennis ball after it has left contact with the racquet and before it touches the ground?

(A) 1 only.
(B) 1 and 2.
(C) 1 and 3.
(D) 2 and 3.
(E) 1, 2, and 3.

Figure 3.10: Question 30 on the FCI

Table 3.13: Breakdown of Question 30 Responses by Student Subset

	Total Responses	Most Common Distractor	A	B	**C**	D	E
Experts	726	E	2%	8%	**18%**	3%	70%
Non-Experts	585	E	3%	12%	**13%**	7%	65%

p-value: 5.21e-5

Preconceptions of both groups:

Students of each group overwhelmingly chose E, not noticing that once the ball has left the racket there is no force from the racket acting upon the ball.

Preconceptions of experts:

Experts' responses peaked sharply at E, with B being the next chosen distractor, ignoring air resistance.

Preconceptions of non-experts:

Non-experts are slightly more varied, but not by much. Many more students chose D than experts, which would imply that their thinking is that since the ball is traveling upward, all the forces must be pushing it up.

CHAPTER 4
DISCUSSION

4.1 Common Preconceptions

Far and away, the most common preconception among all students is the idea that motion *requires* a force. That is, if an object is in motion, a force must be acting on it somehow in the direction of motion. This was constantly popping up among responses from both expert classes and non-expert classes, as seen in Questions 5, 11, 25, and 30. Filtering those parts of the responses out can lead to points of interest for experts and non-experts that can formulate an overall pattern in things that these students tend to miss or over-examine.

4.2 Expert Preconceptions

One interesting thing in five of the nine questions examined within Chapter 3 is that experts seem to cluster their answers much more than their peers. In the majority of questions, expert's responses would peak at one response. Experts also tend to be more homogeneous in their thinking than non-experts. They appear to have a single line of thinking and follow it without much question. This should be looked into more in future work to determine whether it is significant or not.

The majority of preconceptions experts face are shared by non-experts as well, with a few notable exceptions, one being on Question 29, the misapplication of the term "net" force, something a non-expert probably wouldn't have any recognition of.

4.3 Non-Expert Preconceptions

One interesting preconception among non-experts is the seeming idea that *past events don't matter*. On Question 21, non-experts regularly chose the distractor that only had effects from the introduced force, not the prior motion. This could be used to update instruction in classes by emphasizing that an object's entire motion is connected and that forces only act as a change in acceleration - and not necessarily a dramatic one, at that.

Non-experts, as one would expect, also don't fully grasp the sheer number of forces that can act on an object at any given time. How can a student reply that a supporting force from the floor is acting on a chair when they don't know the term for "normal force," as seen in Questions 11 and 29? This is most likely related to the preconception that motion requires force, as it implies that all forces then create motion.

The duration of a force's effects also seems to generate some confusion for non-experts, as shown in Question 30. The amount of force that acts on an object has no bearing on the duration of the force; e.g., just because a force is large doesn't mean an object will accelerate well after it's been acted upon. An object will only be accelerating while a force is present. Looking at the responses for non-experts for Question 30 shows a somewhat larger percentage of responses than contain "a force by the 'hit'" present.

Overall, the main preconceptions that plague non-experts, differing from their more expert-like counterparts, are that past events can be neglected for the current forces, a stabilizing force is still a force, and a force's size has no bearing on the duration of an object's acceleration. All of these can be understood through basic physical interactions. A few that come to mind are pushing a swing - a person on a swing will still experience an accelerating force after a push has been provided (from gravity), but it is easily interpreted as a large push accelerating an object well after the push has subsided. The swing example also serves double-duty here, in that an object on a swing's prior oscillatory motion probably won't be noticed while swinging (e.g., a person swinging faster will require a larger force to continue swinging back.)

These preconceptions can be easily explained by everyday childhood interactions with the world that most students have experienced.

4.4 Post-Test Data

While not investigated in-depth, the post-test data for every year and semester was examined to determine if the preconceptions discussed remain at the end of the semester. The statistical analysis to determine if the responses overall were statistically significant between the two groups and an analysis to determine if the choice of most common distractor was statistically significant were run, and only one question remained between both the pre- and post-test data. This was Question 25, which only utilizes the preconception that friction is entirely dependent upon weight. However, to get a more robust understanding of how these preconceptions change over the duration of a semester, comparisons between pre- and post-test data for each of the questions discussed previously are in Tables 4.1-4.18. Ideally this information will operate as a jumping-off point for future research on preconceptions' effects on teaching, as only broad-strokes analysis has been done.

4.4.1 Question 2

Table 4.1: Breakdown of Question 2 Pre-Test Responses by Student Subset

	Total Responses	Most Common Distractor	**A**	B	C	D	E
Experts	734	D	**28%**	26%	9%	31%	6%
Non-Experts	599	D	**19%**	26%	7%	41%	8%

Table 4.2: Breakdown of Question 2 Post-test Responses by Student Subset

	Total Responses	Most Common Distractor	**A**	B	C	D	E
Experts	545	B	**39%**	28%	8%	22%	3%
Non-Experts	490	D	**19%**	30%	11%	36%	4%

Non-experts don't seem to change much between the pre- and post-test examination. Experts, however, made an increase in correct responses. In both groups, the most common preconceptions consist of a smaller portion of overall responses.

4.4.2 Question 5

Table 4.3: Breakdown of Question 5 Pre-Test Responses by Student Subset

	Total Responses	Most Common Distractor	A	**B**	C	D	E
Experts	736	C	4%	**16%**	32%	26%	22%
Non-Experts	599	C	9%	**11%**	45%	16%	20%

Correct responses improved between the pre- and post-test, showing that learning has occurred. However, the number of students who still selected a response that involves a force in the direction of motion (D) has risen significantly among non-experts.

Table 4.4: Breakdown of Question 5 Post-Test Responses by Student Subset

	Total Responses	Most Common Distractor	A	**B**	C	D	E
Experts	545	D	5%	**31%**	24%	26%	15%
Non-Experts	491	D	8%	**13%**	26%	32%	22%

4.4.3 Question 11

Table 4.5: Breakdown of Question 11 Pre-Test Responses by Student Subset

	Total Responses	Most Common Distractor	A	B	C	**D**	E
Experts	738	C	11%	27%	41%	**17%**	4%
Non-Experts	599	B	8%	41%	38%	**11%**	2%

Table 4.6: Breakdown of Question 11 Post-Test Responses by Student Subset

	Total Responses	Most Common Distractor	A	B	C	**D**	E
Experts	547	C	7%	8%	43%	**39%**	3%
Non-Experts	491	C	7%	20%	55%	**17%**	2%

Question 11 changes include an increase in correct responses but also an increase in selection of C for both groups. C is incorrect because it also includes a force in the direction of motion.

4.4.4 Question 15

Table 4.7: Breakdown of Question 15 Pre-Test Responses by Student Subset

	Total Responses	Most Common Distractor	**A**	B	C	D	E
Experts	735	C	**31%**	10%	53%	6%	1%
Non-Experts	600	C	**23%**	18%	49%	9%	1%

Table 4.8: Breakdown of Question 15 Post-Test Responses by Student Subset

	Total Responses	Most Common Distractor	**A**	B	C	D	E
Experts	547	C	**50%**	7%	40%	2%	1%
Non-Experts	491	C	**29%**	15%	51%	4%	1%

Only slightly more students selected the most common preconception among non-experts in the post-test, with every other incorrect answer receiving fewer responses.

4.4.5 Question 21

Table 4.9: Breakdown of Question 21 Pre-Test Responses by Student Subset

	Total Responses	Most Common Distractor	A	B	C	D	**E**
Experts	735	C	10%	14%	24%	14%	**38%**
Non-Experts	597	B	11%	26%	24%	15%	**24%**

Table 4.10: Breakdown of Question 21 Post-Test Responses by Student Subset

	Total Responses	Most Common Distractor	A	B	C	D	**E**
Experts	545	C	9%	12%	25%	12%	**42%**
Non-Experts	490	C	13%	20%	26%	15%	**27%**

There isn't a lot of change between the pre- and post-test data for Question 21. There is a slight increase in most common distractor responses, as well as an increase in correct responses.

4.4.6 Question 25

Table 4.11: Breakdown of Question 25 Pre-Test Responses by Student Subset

	Total Responses	Most Common Distractor	A	B	**C**	D	E
Experts	734	D	3%	14%	**20%**	48%	15%
Non-Experts	597	D	5%	22%	**15%**	37%	20%

Table 4.12: Breakdown of Question 25 Post-Test Responses by Student Subset

	Total Responses	Most Common Distractor	A	B	**C**	D	E
Experts	545	D	6%	11%	**24%**	48%	12%
Non-Experts	487	D	7%	18%	**18%**	41%	16%

Roughly four percent more non-experts selected the most common distractor in the post-test data than the pre-test, with a slight increase in correct responses.

4.4.7 Question 28

Table 4.13: Breakdown of Question 28 Pre-Test responses by Student Subset

	Total Responses	Most Common Distractor	A	B	C	D	**E**
Experts	727	D	3%	10%	7%	48%	**32%**
Non-Experts	587	D	2%	20%	11%	53%	**14%**

Table 4.14: Breakdown of Question 28 Post-Test Responses by Student Subset

	Total Responses	Most Common Distractor	A	B	C	D	**E**
Experts	539	D	3%	4%	5%	31%	**56%**
Non-Experts	479	D	3%	10%	8%	56%	**24%**

Question 28 had a dramatic increase in correct responses from pre-test to post-test data among both groups, with non-experts also experiencing a slight increase in the most common distractor.

4.4.8 Question 29

Table 4.15: Breakdown of Question 29 Pre-Test Responses by Student Subset

	Total Responses	Most Common Distractor	A	**B**	C	D	E
Experts	729	D	18%	**44%**	2%	32%	4%
Non-Experts	586	A	32%	**32%**	5%	25%	7%

Table 4.16: Breakdown of Question 29 Post-Test Responses by Student Subset

	Total Responses	Most Common Distractor	A	**B**	C	D	E
Experts	538	D	4%	**70%**	3%	21%	2%
Non-Experts	480	D	9%	**58%**	5%	26%	2%

Question 29 had another dramatic increase in correct response for both groups between pre-test and post-test data, with a slight increase in the distractor D for non-experts.

4.4.9 Question 30

Table 4.17: Breakdown of Question 30 Pre-Test Responses by Student Subset

	Total Responses	Most Common Distractor	A	B	**C**	D	E
Experts	726	E	2%	8%	**18%**	3%	70%
Non-Experts	585	E	3%	12%	**13%**	7%	65%

Table 4.18: Breakdown of Question 30 Post-Test Responses by Student Subset

	Total Responses	Most Common Distractor	A	B	**C**	D	E
Experts	539	E	5%	9%	**29%**	4%	52%
Non-Experts	480	E	3%	11%	**18%**	5%	63%

Question 30 had an increase in correct responses between pre- and post-test data, with a slight decrease in most common distractor responses for non-experts and a rather large increase for experts.

4.5 Future Work

If a person who has access to unredacted information is able to, future work would involve seeing how both a student's CLASS responses and FCI responses are linked, rather than the path taken here of generalizing per course. This will give much more precision in any sort of analysis and a continuation of what has been done so far.

An analysis of the groupings of answers could also be worth looking into. Experts seem to cluster together, implying a high level of confidence in their responses. This could be done by looking at the responses of non-experts and experts on the CLASS post-test.

While somewhat unrelated to the original research question, any study pertaining to gender differences would be considered important, as there is a notably large difference in gender distribution within physics and engineering. It's also already fairly well documented that men tend to overestimate their competence while women tend to underestimate theirs, which might have an effect on preconceptions [20].

Further analysis on the differences between pre-test and post-test data should also be considered. While definite conclusions on this change is largely left untouched, it would serve as a good starting point to see the effect teaching has on students' preconceptions.

4.6 Conclusions

In conclusion, there are a few general differences in preconceptions between students within PHYS 253, generally considered more expert-like, and students within PHYS 150 and PHYS 210, generally considered less expert-like, as defined by the CLASS. These preconceptions tend towards non-experts not taking into account an object's past trajectory when figuring out future trajectories, an interesting interpretation about motion requiring force, and that the size of a force acting upon an object does not have any bearing on the duration of an object's acceleration. For experts, the most notable preconception was misapplications of the definition of "net" force.

Both subsets of students had the common preconception that motion *requires* force. Upon brief analysis of the changes between the pre-test and post-test data, there is a lack of statistical significance in responses between the two groups. However, with non-experts it seems that responses are largely split between the correct response and the same (or similar) distractor ,as in the pre-test data. This can be seen in Questions 28, 25, 15, 11, and 5. While there is no longer a significant change between the two groups and their most common distractor in the post-test data, it does appear as though the common preconceptions between the two groups remains. However, without further analysis, this is mere speculation. It is hoped that this information will be used to more directly combat students' preconceptions, which have been shown to correlate to increased scores on the FCI [3][15]. Notably, including rigorous mathematical definitions and connections to these preconceptions will hopefully remove them from students' minds.

Overall, it seems as though the majority of preconceptions that experts selected were also selected by their non-expert counterparts. This implies that experts do not make unique errors based on their prior knowledge. This research has shown, however, that non-experts do hold considerably different preconceptions than their expert-like peers, most notably that if they believe motion requires a force, all forces must be in the direction of an object's motion. So while the initial hypothesis has not been proven true, some glimmer of knowledge has still been obtained.

REFERENCES

[1] John Clement. "Students' preconceptions in introductory mechanics". In: *American Journal of Physics* 50.1 (1982), pp. 66–71. URL: https://aapt.scitation.org/doi/abs/10.1119/1.12989.

[2] Ibrahim Abou Halloun and David Hestenes. "Common sense concepts about motion". In: *American Journal of Physics* 53.11 (1985), pp. 1056–1065. URL: https://aapt.scitation.org/doi/abs/10.1119/1.14031.

[3] Paula Heron. "Identifying and Addressing Difficulties: Reflections on the empirical and theoretical basis of an influential approach to improving physics education". In: *Getting Started in PER*. 4th ed. Vol. 2. 21. URL: https://www.compadre.org/Repository/document/ServeFile.cfm?ID=14724&DocID=4885.

[4] Leah Savion. "Clinging to Discredited Beliefs: The Larger Cognitive Story". In: *Journal of the Scholarship of Teaching and Learning* 9.1 (2012), pp. 81–92.

[5] George J. Posner et al. "Accommodation of a scientific conception: Toward a theory of conceptual change". In: *Science Education* 66.2 (1982), pp. 211–227. DOI: 10.1002/sce.3730660207. URL: https://doi.org/10.1002/sce.3730660207.

[6] Michelene T. H. Chi. "Commonsense Conceptions of Emergent Processes: Why Some Misconceptions Are Robust". In: *Journal of the Learning Sciences* 14.2 (Apr. 2005), pp. 161–199. DOI: 10.1207/s15327809jls1402{_}1. URL: https://doi.org/10.1207/s15327809jls1402_1.

[7] J. Kruger and D. Dunning. "Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments." In: *Journal of Personality and Social Psychology* 77.6 (1999), pp. 1121–1134.

[8] David Hestenes, Malcolm Wells, and Gregg Swackhamer. "Force concept inventory". In: *The Physics Teacher* 30.3 (1992), pp. 141–158. DOI: 10.1119/1.2343497. URL: https://doi.org/10.1119/1.2343497.

[9] W. K. Adams et al. "New instrument for measuring student beliefs about physics and learning physics: The Colorado Learning Attitudes about Science Survey". In: *Physical Review Special Topics - Physics Education Research* 2.1 (Jan. 2006), pp. 010101–. DOI: 10.1103/PhysRevSTPER.2.010101. URL: https://link.aps.org/doi/10.1103/PhysRevSTPER.2.010101.

[10] The Editors of Encyclopaedia Britannica. *Spontaneous generation.*

[11] N. Sanjay Rebello and Dean A. Zollman. "The effect of distracters on student performance on the force concept inventory". In: *American Journal of Physics* 72.1 (2003), pp. 116–125. DOI: 10.1119/1.1629091. URL: https://doi.org/10.1119/1.1629091.

[12] 2019. URL: https://www.surveymonkey.com/mp/likert-scale/.

[13] W. K. Adams et al. "The Design and Validation of the Colorado Learning Attitudes about Science Survey". In: *AIP Conference Proceedings* 790.1 (2005), pp. 45–48. DOI: 10.1063/1.2084697. URL: https://aip.scitation.org/doi/abs/10.1063/1.2084697.

[14] Richard R. Hake. "Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses". In: *American Journal of Physics* 66.1 (1998), pp. 64–74. DOI: 10.1119/1.18809. URL: https://doi.org/10.1119/1.18809.

[15] Antti Savinainen and Philip Scott. "Using the Force Concept Inventory to monitor student learning and to plan teaching". In: 37.1 (2002), pp. 53–58. DOI: 10.1088/0031-9120/37/1/307. URL: http://dx.doi.org/10.1088/0031-9120/37/1/307.

[16] URL: https://learningassistantalliance.org/.

[17] URL: https://www.physport.org/assessments/guides/Open.cfm?G=3&R=160.

[18] 2018. URL: http://catalog.niu.edu/.

[19] Ronald A. Fisher. *Statistical Methods for Research Workers.*

[20] Chiungjung Huang. "Gender differences in academic self-efficacy: a meta-analysis". In: *European Journal of Psychology of Education* 28.1 (2013), pp. 1–35. DOI: 10.1007/s10212-011-0097-y. URL: https://doi.org/10.1007/s10212-011-0097-y.

[21] URL: https://newonlinecourses.science.psu.edu/stat500/node/56/.

APPENDIX A

P-VALUE DATA

In Table A.1, the p-values are for the significance of differing frequencies in total distractor responses between experts and non-experts. In Table A.2, the p-values are for the significance of different frequencies of most common distractor between experts and non-experts. Statistically significant responses in Table A.1 are in bold.

Question	p-value
1	0.385
2	**0.029**
3	0.6
4	**0.045**
5	**2.8e-9**
6	0.118
7	0.532
8	0.727
9	0.675
10	0.717
11	**9.38e-5**
12	**0.023**
13	0.056
14	0.628
15	**2.81e-4**
16	0.866
17	0.0513
18	**0.011**
19	0.479
20	0.793
21	**0.002**
22	0.677
23	0.092
24	0.199
25	**9.19e-7**
26	0.118
27	0.741
28	**0.001**
29	**4.71e-9**
30	**1.91e-4**

Table A.1: p-values of Total Responses

Question	p-value
2	0.0156
5	1.51e-4
11	0.0159
15	2.778e-5
21	0.04
25	6.45e-8
28	2.95e-3
29	6.058e-10
30	5.21e-5

Table A.2: p-values of Most Common Distractor from Statistically Significant Questions

APPENDIX B

DATA BY YEAR

In Figures B.1 and B.2, the “other” STEM majors involve majors like biology and math. PHYS 210 is one of the suggested required courses for biology majors at Northern Illinois University.

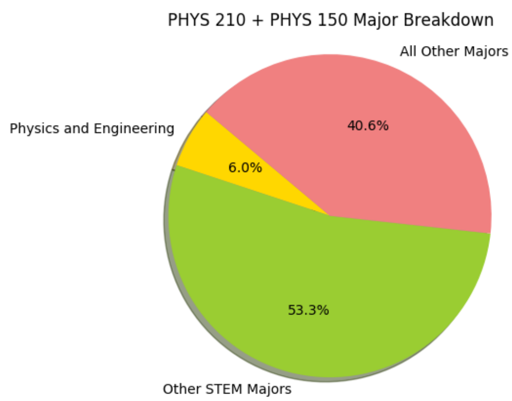


Figure B.1: PHYS 210 and PHYS 150 Breakdown by Major for Fall 2018.

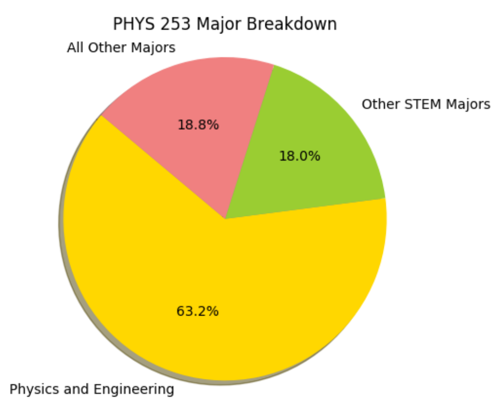


Figure B.2: PHYS 253 Breakdown by Major for Fall 2018.

Table B.1: Spring 2015 Expert FCI Pre-Test

Question	Percentage of Right Answers	Total Responses	Common Distractor	A	B	C	D	E
1	67%	135	4	8%	7%	67%	14%	3%
2	31%	134	4	31%	19%	10%	33%	7%
3	40%	134	1	30%	12%	40%	6%	13%
4	31%	135	1	68%	1	0%	0%	31%
5	15%	136	5	4%	15%	24%	28%	29%
6	82%	136	1	14%	82%	3%	1	1
7	68%	136	5	10%	68%	6%	3%	12%
8	48%	135	1	30%	48%	1	5%	15%
9	37%	134	2	4%	30%	24%	6%	37%
10	50%	136	4	50%	2%	14%	17%	17%
11	15%	136	3	11	26%	43%	15%	5%
12	82%	136	3	1%	82%	16%	1%	1%
13	13	136	3	8%	28%	51%	13	0%
14	42%	136	1	28%	18%	12%	42%	0%
15	25%	134	3	25%	9%	60%	5%	1%
16	56%	134	3	56%	5%	27%	7%	4%
17	7%	136	1	63%	7%	3%	19%	8%
18	11%	136	5	1%	11%	10%	36%	43%
19	42%	136	4	15%	7%	8%	29%	42%
20	33%	135	3	21%	3	36%	33%	7%
21	35%	136	3	9%	17%	29%	11%	35%
22	39%	136	4	25%	39%	2%	30%	4
23	46%	136	3	8%	46%	26%	12%	8%
24	70%	136	3	70%	2%	15%	5%	7%
25	16%	135	4	4	13%	16%	49%	19%
26	5%	136	1	42%	32%	3%	18%	5%
27	65%	136	1	21%	12%	65%	2%	1
28	37%	136	4	4	9%	6%	45%	37%
29	46%	136	4	21%	46%	0%	32%	1%
30	14%	136	5	2%	4%	14%	4%	75%

Table B.2: Fall 2015 Expert FCI Pre-Test

Question	Percentage of Right Answers	Total Responses	Common Distractor	A	B	C	D	E
1	64%	209	4	13%	7%	64%	14%	2%
2	28%	210	4	28%	25%	8%	30%	9%
3	48%	210	1	28%	15%	48%	4%	6%
4	32%	210	1	63%	4	0%	1	32%
5	15%	210	3	5	15%	38%	23%	19%
6	78%	210	1	18%	78%	3%	0%	0%
7	64%	209	5	11%	64%	9%	5	11%
8	50%	210	1	21%	50%	1	9%	18%
9	44%	210	2	5%	28%	18%	5%	44%
10	54%	210	4	54%	5%	12%	20%	9%
11	19%	210	3	10%	26%	41%	19%	4%
12	80%	210	3	1%	80%	17%	2%	1%
13	19%	210	3	12%	24%	44%	19%	0%
14	48%	210	1	25%	17%	11%	48%	0%
15	32%	210	3	32%	9%	55%	4%	0%
16	60%	210	3	60%	6%	27%	6%	1%
17	10%	210	1	66%	10%	1	17	6%
18	18	210	4	1%	18	18	34%	30%
19	48%	210	4	11%	5%	5%	31%	48%
20	33%	209	3	24%	8%	30%	33%	5%
21	39%	209	3	10%	12%	22%	17%	39%
22	44%	210	1	25%	44%	5%	23%	2%
23	47%	210	3	4%	47%	22%	19%	8%
24	70%	210	3	70%	6%	13%	4%	7%
25	21%	210	4	2%	15%	21%	50%	11%
26	8%	210	1	41%	36%	3%	12%	8%
27	64%	210	1	21%	13%	64%	1	0%
28	33%	207	4	1%	6%	8%	51%	33%
29	46%	207	4	13%	46%	1%	34%	5%
30	19%	207	5	0%	5	19%	2%	73%

Table B.3: Fall 2015 Non-Expert FCI Pre-Test

Question	Percentage of Right Answers	Total Responses	Common Distractor	A	B	C	D	E
1	41%	137	4	18%	7%	41%	31%	3%
2	21%	137	4	21%	22%	5%	45%	7%
3	27%	137	1	34%	21%	27%	8%	9%
4	15%	137	1	80%	1	1	3%	15%
5	9%	136	3	7%	9%	50%	12%	22%
6	66%	137	1	20%	66%	10%	2%	1
7	49%	137	1	17%	49%	15%	9%	11%
8	39%	137	1	29%	39%	2%	9%	20%
9	29%	137	2	4%	32%	23%	11%	29%
10	35%	137	4	35%	1%	23%	30%	12%
11	8%	137	2	11	42%	36%	8%	3%
12	57%	137	3	1%	57%	33%	6%	4%
13	8%	137	3	8%	34%	50%	8%	0%
14	23%	137	1	37%	20%	19%	23%	0%
15	18%	137	3	18%	14%	57%	10%	1%
16	43%	137	3	43%	8%	37%	9%	2%
17	4%	137	1	68%	4%	1	21%	6%
18	4	137	4	4	4	24%	39%	28%
19	39%	137	4	10%	4	9%	38%	39%
20	16%	137	3	31%	7%	36%	16%	10%
21	23%	137	2	10%	27%	24%	15%	23%
22	30%	137	1	31%	30%	6%	29%	4%
23	26%	137	3	13%	26%	32%	22%	7%
24	57%	136	3	57%	1%	28%	1%	12%
25	12%	137	4	3%	18%	12%	42%	26%
26	1	137	1	49%	33%	4%	13%	1
27	49%	135	1	39%	12%	49%	0%	1
28	11%	134	4	2%	22%	7%	57%	11%
29	23%	134	1	37%	23%	1	31%	8%
30	11%	134	5	1%	8%	11%	8%	71%

Table B.4: Spring 2016 Expert FCI Pre-Test

Question	Percentage of Right Answers	Total Responses	Common Distractor	A	B	C	D	E
1	67%	119	1	13%	8%	67%	9%	3%
2	28%	119	2	28%	33%	9%	27%	3%
3	49%	119	1	24%	18%	49%	7%	3
4	41%	119	1	55%	2%	2%	0%	41%
5	16%	118	4	2%	16%	27%	34%	21%
6	78%	119	1	17%	78%	4%	1	0%
7	61%	119	1	13%	61%	8%	8%	8%
8	47%	119	1	23%	47%	1	12%	18%
9	36%	119	2	8%	26%	20%	10%	36%
10	56%	119	4	56%	3%	9%	20%	11%
11	17%	119	3	8%	25%	49%	17%	2%
12	76%	119	3	3	76%	18%	2%	1%
13	17%	119	3	9%	26%	46%	17%	2%
14	46%	119	1	30%	15%	8%	46%	0%
15	39%	119	3	39%	9%	49%	3	0%
16	68%	119	3	68%	5%	23%	3	1%
17	10%	118	1	64%	10%	2%	17	7%
18	17%	119	4	0%	17%	16%	35%	32%
19	45%	119	4	9%	4	9%	32%	45%
20	38%	119	3	14%	10%	30%	38%	8%
21	40%	119	3	8%	13%	29%	11%	40%
22	34%	119	4	28%	34%	3%	32%	3%
23	44%	118	3	5%	44%	21%	21%	8%
24	69%	119	3	69%	4%	19%	2%	6%
25	13%	119	4	2%	13%	13%	58%	13%
26	4%	119	1	40%	39%	4%	13%	4%
27	67%	119	1	19%	12%	67%	2%	0%
28	35%	118	4	3%	7%	8%	47%	35%
29	46%	118	4	16%	46%	3%	31%	4
30	21%	118	5	1%	8%	21%	5	64%

Table B.5: Spring 2016 Non-Expert FCI Pre-Test

Question	Percentage of Right Answers	Total Responses	Common Distractor	A	B	C	D	E
1	51%	141	4	17%	8%	51%	20%	4
2	19%	141	4	19%	29%	5%	38%	9%
3	31%	141	1	30%	21%	31%	5%	13%
4	43%	141	1	55%	1	1	0%	43%
5	11%	141	3	11%	11%	39%	18%	21%
6	70%	141	1	20%	70%	8%	1	1
7	49%	141	1	16%	49%	12%	9%	15%
8	32%	141	1	30%	32%	0%	13%	25%
9	26%	141	2	7%	35%	25%	6%	26%
10	35%	141	4	35%	5%	14%	26%	20%
11	13%	141	3	6%	39%	41%	13%	1%
12	53%	141	3	1%	53%	33%	9%	4%
13	7%	141	3	10%	35%	48%	7%	1%
14	21%	141	1	43%	23%	13%	21%	1
15	38%	141	3	38%	16%	40%	6%	0%
16	48%	141	3	48%	6%	33%	6%	6%
17	13%	141	1	57%	13%	3%	14%	13%
18	7%	141	4	1%	7%	22%	38%	31%
19	25%	141	4	16%	5%	3%	52%	25%
20	14%	140	3	24%	5%	55%	14%	3
21	26%	139	3	14%	21	25%	14%	26%
22	33%	138	4	26%	33%	4	34%	3%
23	31%	139	3	13%	31%	27%	25%	4%
24	48%	139	3	48%	4%	30%	6%	13%
25	14%	139	4	6%	19%	14%	40%	19%
26	3%	139	1	40%	32%	9%	17%	3%
27	45%	139	1	37%	15%	45%	3%	0%
28	23%	135	4	2%	13%	12%	50%	23%
29	43%	134	1	24%	43%	8%	20%	5%
30	13%	135	5	1%	9%	13%	6%	71%

Table B.6: Fall 2016 Expert FCI Pre-Test

Question	Percentage of Right Answers	Total Responses	Common Distractor	A	B	C	D	E
1	58%	195	4	12%	11%	58%	17%	2%
2	24%	194	4	24%	28%	9%	34%	5%
3	42%	195	2	24%	25%	42%	1%	9%
4	25%	195	1	72%	3%	1	0%	25%
5	19%	194	3	2%	19%	34%	25%	21%
6	71%	195	1	23%	71%	4%	2%	1
7	56%	195	1	16%	56%	8%	7	12%
8	48%	195	1	19%	48%	2%	11%	19%
9	31%	193	2	8%	26%	24%	11%	31%
10	51%	195	4	51%	4	14%	22%	10
11	17%	195	3	13%	31%	36%	17%	3
12	71%	195	3	2%	71%	22%	5%	1%
13	16%	195	3	10%	26%	47%	16%	1%
14	47%	194	1	26%	15%	11%	47%	1
15	29%	195	3	29%	9%	51%	10%	2%
16	54%	195	3	54%	8%	26%	8%	4%
17	10%	194	1	62%	10%	3%	14%	10%
18	18	195	4	4	18	20%	32%	27%
19	42%	195	4	9%	4	9%	36%	42%
20	27%	195	3	24%	6%	39%	27%	5%
21	38%	194	3	13%	15%	21	12%	38%
22	42%	195	1	24%	42%	8%	24%	4%
23	39%	195	3	9%	39%	24%	22%	6%
24	61%	195	3	61%	4%	25%	6%	5%
25	21%	194	4	6%	15%	21%	40%	17%
26	7%	195	1	39%	29%	7%	18%	7%
27	62%	195	1	24%	12%	62%	3%	0%
28	26%	191	4	4	15%	8%	46%	26%
29	42%	192	4	21%	42%	3%	32%	3%
30	17%	190	5	4%	11%	17%	2%	66%

Table B.7: Fall 2016 Non-Expert FCI Pre-Test

Question	Percentage of Right Answers	Total Responses	Common Distractor	A	B	C	D	E
1	47%	181	4	15%	12%	47%	20%	5%
2	22%	180	4	22%	26%	7%	42%	3%
3	26%	181	1	31%	25%	26%	5%	13%
4	14%	181	1	80%	3%	1	2%	14%
5	14%	181	3	10%	14%	43%	13%	20%
6	64%	180	1	24%	64%	7%	2%	2%
7	49%	180	5	17%	49%	6%	9%	18%
8	36%	181	1	29%	36%	2%	11%	22%
9	23%	179	2	7%	34%	28%	8%	23%
10	41%	181	4	41%	8%	14%	25%	12%
11	12%	180	2	9%	43%	34%	12%	2
12	49%	180	3	1%	49%	38%	7%	4%
13	10%	180	3	17%	31%	40%	10%	2%
14	24%	180	1	38%	26%	11%	24%	0%
15	17%	181	3	17%	22%	51%	9%	1%
16	39%	181	3	39%	10%	39%	7%	5%
17	13%	180	1	62%	13%	4%	13%	8%
18	11%	180	4	2%	11%	18	36%	34%
19	30%	180	4	12%	6%	8%	44%	30%
20	16%	179	3	27%	11%	41%	16%	6%
21	24%	180	2	11%	25%	22%	18%	24%
22	40%	180	4	23%	40%	9%	24%	4
23	23	180	4	17%	23	23	29%	8%
24	52%	180	3	52%	6%	27%	5%	10%
25	19%	180	4	5%	24%	19%	34%	17%
26	4%	181	1	42%	31%	9%	14%	4%
27	46%	181	1	30%	19%	46%	4%	1
28	11%	179	4	3%	21%	10%	55%	11%
29	28%	179	1	34%	28%	7%	25%	7%
30	14%	178	5	3%	19%	14%	7%	57%

Table B.8: Spring 2017 Expert FCI Pre-Test

Question	Percentage of Right Answers	Total Responses	Common Distractor	A	B	C	D	E
1	62%	78	1	13%	10%	62%	10%	5%
2	32%	77	2	32%	27%	10%	26%	4%
3	31%	77	1	31%	25%	31%	3	10%
4	22%	77	1	73%	3%	0%	3%	22%
5	18%	78	3	4%	18%	36%	23%	19%
6	74%	77	1	17%	74%	5%	4%	0%
7	57%	77	1	14%	57%	13%	3%	13%
8	45%	77	1	22%	45%	0%	16%	17%
9	26%	77	2	3%	34%	32%	5%	26%
10	58%	78	4	58%	3%	10	21%	9%
11	18%	78	3	13%	28%	37%	18%	4%
12	70%	77	3	0%	70%	25%	3	3
13	19%	77	3	9%	26%	45%	19%	0%
14	44%	77	1	31%	8%	17%	44%	0%
15	30%	77	3	30%	16%	49%	4%	1%
16	56%	77	3	56%	12%	29%	3	1%
17	10%	77	1	70%	10%	1	12%	6%
18	10%	77	4	3%	10%	13%	43%	31%
19	40%	77	4	10%	3%	5%	42%	40%
20	25%	76	3	16%	14%	39%	25%	5%
21	38%	77	3	12%	13%	22%	16%	38%
22	42%	76	1	26%	42%	4%	22	5%
23	37%	76	4	16%	37%	21%	25%	1%
24	64%	77	3	64%	4%	22%	4%	6%
25	34%	76	4	0%	7%	34%	46%	13%
26	9%	76	1	42%	39%	3%	7%	9%
27	65%	77	1	22%	12%	65%	0%	1
28	28	75	4	0%	15%	4	53%	28
29	42%	76	4	21%	42%	3%	32%	3%
30	16%	75	5	0%	12%	16%	3%	69%

Table B.9: Spring 2017 Non-Expert FCI Pre-Test

Question	Percentage of Right Answers	Total Responses	Common Distractor	A	B	C	D	E
1	47%	141	4	15%	13%	47%	16%	9%
2	11%	141	4	11%	27%	11%	39%	12%
3	26%	141	1	34%	23%	26%	8%	10%
4	14%	141	1	77%	3%	1	5%	14%
5	9%	141	3	6%	9%	48%	19%	18%
6	65%	141	1	24%	65%	8%	2%	1
7	53%	141	1	13%	53%	13%	11%	9%
8	39%	141	1	31%	39%	2%	9%	18%
9	26%	141	2	11%	38%	18%	6%	26%
10	26%	140	4	26%	6%	11%	41%	17%
11	10%	141	3	7%	39%	40%	10%	4%
12	52%	141	3	3	52%	31%	8%	6%
13	7%	141	3	16%	35%	41%	7%	1%
14	20%	141	1	44%	23%	13%	20%	0%
15	21%	141	3	21%	18%	49%	9%	3
16	39%	141	3	39%	11%	35%	11%	4%
17	11%	141	1	55%	11%	9%	13%	12%
18	11%	141	3	6%	11%	32%	25%	27%
19	25%	141	4	16%	6%	9%	45%	25%
20	12%	141	3	33%	10%	36%	12%	9%
21	23%	141	2	9%	30%	26%	12%	23%
22	33%	141	4	26%	33%	7%	29%	5%
23	26%	141	3	12%	26%	33%	22%	6%
24	48%	141	3	48%	6%	30%	5%	11%
25	15%	141	4	6%	28%	15%	33%	18%
26	4%	141	1	40%	35%	8%	14%	4%
27	43%	141	1	34%	20%	43%	2%	1
28	10%	139	4	0%	23%	17%	50%	10%
29	35%	139	1	31%	35%	5%	23%	6%
30	14%	138	5	4%	9%	14%	8%	65%

Table B.10: All Experts FCI Pre-Test

Question	Percentage of Right Answers	Total Responses	Common Distractor	A	B	C	D	E
1	63%	736	4	12%	9%	63%	14%	3%
2	28%	734	4	28%	26%	9%	31%	6%
3	43%	735	1	27%	18%	43%	4%	8%
4	30%	736	1	66%	2%	1	1	30%
5	16%	736	3	4%	16%	32%	26%	22%
6	77%	737	1	18%	77%	4%	1	1
7	62%	736	1	13%	62%	9%	6%	11%
8	48%	736	1	23%	48%	1	10%	18%
9	36%	733	2	6%	28%	22%	8%	36%
10	53%	738	4	53%	4	12%	20%	11%
11	17%	738	3	11	27%	41%	17%	4%
12	76%	737	3	1%	76%	19%	2%	1%
13	17%	737	3	10%	26%	47%	17%	1%
14	46%	736	1	27%	15%	12%	46%	0%
15	31%	735	3	31%	10%	53%	6%	1%
16	58%	735	3	58%	7%	26%	6%	3
17	10%	735	1	65%	10%	2%	16%	8%
18	16%	737	4	2%	16%	16%	35%	32%
19	44%	737	4	11%	5%	7%	33%	44%
20	31%	734	3	21%	7%	35%	31%	6%
21	38%	735	3	10%	14%	24%	14%	38%
22	41%	736	4	25%	41%	5%	26%	3%
23	43%	735	3	8%	43%	23	19%	7%
24	67%	737	3	67%	4%	19%	4%	6%
25	20%	734	4	3%	14%	20%	48%	15%
26	7%	736	1	41%	34%	4%	15%	7%
27	64%	737	1	22%	12%	64%	2%	0%
28	32%	727	4	3%	10%	7%	48%	32%
29	44%	729	4	18%	44%	2%	32%	4
30	18%	726	5	2%	8%	18%	3%	70%

Table B.11: All Non-Experts FCI Pre-Test

Question	Percentage of Right Answers	Total Responses	Common Distractor	A	B	C	D	E
1	46%	600	4	16%	10%	46%	22%	5%
2	19%	599	4	19%	26%	7%	41%	8%
3	27%	600	1	32%	23%	27%	6%	12%
4	21%	600	1	73%	2%	1	2%	21%
5	11%	599	3	9%	11%	45%	16%	20%
6	66%	599	1	22%	66%	8%	2%	2%
7	50%	599	1	16%	50%	11%	10%	14%
8	36%	600	1	30%	36%	2%	11%	21%
9	26%	598	2	7%	35%	24%	8%	26%
10	35%	599	4	35%	5%	15%	30%	15%
11	11	599	2	8%	41%	38%	11	2
12	53%	599	3	2%	53%	34%	7%	5%
13	8%	599	3	13	34%	44%	8%	1%
14	22%	599	1	40%	23%	14	22%	0%
15	23%	600	3	23%	18%	49%	8%	1%
16	42%	600	3	42%	9%	36%	8%	4%
17	11%	599	1	60%	11%	4%	15%	10%
18	8%	599	4	3%	8%	24%	34%	31%
19	30%	599	4	13%	5%	7%	45%	30%
20	14%	597	3	28%	8%	42%	14%	7%
21	24%	597	2	11%	26%	24%	15%	24%
22	34%	596	4	26%	34%	7%	29%	4
23	26%	597	3	14%	26%	29%	25%	6%
24	51%	596	3	51%	5%	29%	4%	11%
25	15%	597	4	5%	22%	15%	37%	20%
26	3%	598	1	43%	32%	8%	14%	3%
27	46%	596	1	34%	17%	46%	3%	1
28	14%	587	4	2%	20%	11%	53%	14%
29	32%	586	1	32%	32%	5%	25%	7%
30	13%	585	5	3%	12%	13%	7%	65%

APPENDIX C

χ^2 TEST OF INDEPENDENCE

The motivation for the χ^2 test of ndependence is to determine if two categorical variables are related. Until there is evidence to suggest they are, the base assumption is that they are not. These are our null and alternative hypotheses. In other words:

- H_0: In the population, the two categorical variables are independent.
- H_1: In the population, the two categorical variables are dependent.

In this study, our two categorical variables are "Experts" and "Non-Experts." The null hypothesis refers to the two variables being unrelated; while the alternative hypothesis refers to some relationship between the variables and the data provided.

Once all data is gathered, it is summarized in a table, referred to as a contingency table or an observed counts table. The first step in answering our question ("Are these variables dependent?") is to first answer the question: "What would our data look like if the two variables are not related?" This requires finding the expected table. This is done for each table cell by multiplying the row total and the column total, then dividing by the sample size. On our Question 2, this looks like Tables C.1 and C.2.

Q2	Distractor 1	Distractor 2	Distractor 3	Distractor 4	Total
Experts	193	67	226	44	530
Non-Experts	155	42	246	45	488
Total	348	109	472	89	1,018

Table C.1: Contingency Table Used for χ^2 Test for Independence

Q2	Distractor 1	Distractor 2	Distractor 3	Distractor 4	Total
Experts	$\frac{530*348}{1,018} = 181.18$	$\frac{530*109}{1,018} = 56.75$	$\frac{530*472}{1,018} = 245.74$	$\frac{530*89}{1,018} = 46.34$	530
Non-Experts	$\frac{488*348}{1,018} = 166.82$	$\frac{488*109}{1,018} = 52.25$	$\frac{488*472}{1,018} = 226.26$	$\frac{488*89}{1,018} = 42.66$	488
Total	348	109	472	89	1,018

Table C.2: Expected Table Used for χ^2 Test for Independence

This expected counts table is supposed to represent what would be expected if the two variables are independent of each other. That is, without a link between being an expert and choosing a certain distractor, what would we expect experts and non-experts to choose? Those values are in the expected table.

The question next becomes, "Are the observed counts so different from the expected counts that we can conclude a relationship between the two variables?" This is where the χ^2 test comes into play. In it we compare each cell's observed count to its respective expected count, calculated as follows:

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

With the degrees of freedom = (rows-1)(columns-1), in our case this is 3. The χ^2 value for Question 2 is:

$$\begin{aligned}\chi^2 = & \frac{(193-181.18)^2}{181.18} + \frac{(67-56.75)^2}{56.75} + \frac{(226-245.74)^2}{245.74} + \frac{(44-46.34)^2}{46.34} \\ & + \frac{(155-166.82)^2}{166.82} + \frac{(42-52.25)^2}{52.25} + \frac{(246-226.26)^2}{226.26} + \frac{(45-42.66)^2}{42.66} \\ = & \ 9.0251\end{aligned}$$

A p-value is then calculated to find the probability of a more "extreme" statistic, using the degrees of freedom, and chances that a more unlikely χ^2 value will be found. In this study, this was done automatically through Python code. If this p-value is less than 0.05 (representing a less than 5% chance the difference between expected and observed values is because of random chance), it is generally considered to be proof of some sort of relationship between the groups and the observables. All this information was found in [21].