ABSTRACT

MEASURING THE IMPACT OF CONTENT UPDATES IN AN UNDERGRADUATE NON-MAJOR’S PHYSICS LABORATORY

by William Patrick Boyle IV

Physics laboratories are a staple of any physics education. They exist to expose students to the material world of physics so that they may have a chance to cement their understanding of abstract ideas. Often these laboratories have students confirming the value of different physical constants or performing experiments to show that a physical law will hold true. While ubiquitous, there is little evidence that demonstrates the effectiveness of the physics laboratory. Most research on the matter highlights the ineffectiveness of such courses. Research groups have suggested different methods to improve the effect of laboratory classes; many focus on improving the conceptual understanding gained from the course. This experiment aims to measure the conceptual understanding and scientific reasoning ability of students in a non-major’s physics laboratory and then update the course to be in closer accord with current research based designs and then measure students again. In this way, any change in conceptual understanding or scientific reasoning ability could be claimed to be a result of the updated laboratory material. After running the experiment, it was found that there was a small measurable difference in conceptual understanding between students who completed the reformed laboratory when compared to the classic laboratory.
MEASURING THE IMPACT OF CONTENT UPDATES IN AN UNDERGRADUATE NON-MAJOR’S PHYSICS LABORATORY

A Thesis

Submitted to the
Faculty of Miami University
in partial fulfillment of
the requirements for the degree of
Master of Science
Department of Physics

by
William Patrick Boyle IV
Miami University
Oxford, Ohio
2017

Advisor: Dr. Jennifer Blue
Reader: Dr. Stephen Alexander
Reader: Dr. Karthik Vishwanath

©2017 William Patrick Boyle IV
This Thesis titled

MEASURING THE IMPACT OF CONTENT UPDATES IN AN UNDERGRADUATE NON-MAJOR’S PHYSICS LABORATORY

by

William Patrick Boyle IV

has been approved for publication by

The College of Arts and Science

and

Department of Physics

____________________________________________________

Dr. Jennifer Blue

____________________________________________________

Dr. Stephen Alexander

____________________________________________________

Dr. Karthik Vishwanath
# Table of Contents

List of Tables ........................................................................................................ iv  
List of Figures ........................................................................................................ v  
Introduction ........................................................................................................... 1  
Literature Review ................................................................................................. 3  
Methods ............................................................................................................... 9  
  Data Collection Design ....................................................................................... 9  
  Post-Lab Addition ............................................................................................... 10  
  Content Modifications ......................................................................................... 11  
Section Consistency .............................................................................................. 13  
Analysis ............................................................................................................... 13  
Institutional Review Board Application ................................................................. 14  
Data ................................................................................................................... 15  
  Demographics .................................................................................................... 15  
  Post-labs ........................................................................................................... 18  
  Reading a Wright Map ......................................................................................... 19  
  Lawson Classroom Test of Scientific Reasoning ................................................. 38  
Discussion .......................................................................................................... 40  
  Results ............................................................................................................... 40  
  Limitations ......................................................................................................... 41  
  Future Works ..................................................................................................... 42  
Bibliography ......................................................................................................... 44  
Appendix A: Questions from post-labs ................................................................. 46  
  Measurement and Uncertainty ........................................................................... 46  
  Standing Waves ................................................................................................. 48  
  Simple Harmonic Motion .................................................................................. 50  
  Generators ......................................................................................................... 51  
  Geometric Optics ............................................................................................... 52  
  Physical Optics .................................................................................................. 54  
  Atomic Spectra .................................................................................................. 56  
  Photoelectric Effect ........................................................................................... 57  
  Absolute Zero .................................................................................................... 59  
Appendix B: Demographic questions .................................................................. 61  
Appendix C: Comparison of pre-lab forms .......................................................... 63  
Appendix D: Comparison of lab assignments ....................................................... 65  
Appendix E: IRB Application .............................................................................. 71  
  Consent Form ..................................................................................................... 77
List of Tables

Table 1: Post-test comparison p-values ........................................................................37
Table 2: Summary of Lawson results ........................................................................39
## List of Figures

- Figure 1: Fall semester demographic data, N=96 ................................................................. 16
- Figure 2: Spring semester demographic data, N = 94 .......................................................... 17
- Figure 3: Measurement and Uncertainty Wright maps ............................................................. 20
- Figure 4: Measurement and Uncertainty post lab scores ......................................................... 21
- Figure 5: Standing Waves Wright maps .................................................................................. 22
- Figure 6: Standing Waves post-lab scores .............................................................................. 23
- Figure 7: Simple Harmonic Motion Wright maps ................................................................. 24
- Figure 8: Simple Harmonic Motion post-lab scores ............................................................... 25
- Figure 9: Generators Wright maps ....................................................................................... 26
- Figure 10: Generators post-lab scores .................................................................................... 27
- Figure 11: Geometric Optics Wright maps ........................................................................... 28
- Figure 12: Geometric Optics post-lab scores ....................................................................... 29
- Figure 13: Physical Optics Wright maps ................................................................................. 30
- Figure 14: Physical Optics post-lab scores ............................................................................ 31
- Figure 15: Atomic Spectra Wright maps ............................................................................... 32
- Figure 16: Atomic Spectra post-lab scores ........................................................................... 33
- Figure 17: Photoelectric Effect Wright maps ........................................................................ 34
- Figure 18: Photoelectric Effect post-lab scores .................................................................... 35
- Figure 19: Absolute Zero Wright maps .................................................................................. 36
- Figure 20: Absolute Zero post-lab scores .............................................................................. 37
- Figure 21: Fall Lawson Test results ...................................................................................... 38
- Figure 22: Spring Lawson Test results ................................................................................... 39
INTRODUCTION

This experiment is an attempt to measure the effects of changes made to a laboratory class based on current research. Laboratory activities, or labs, occupy a special place in physics education; many professors feel that they are necessary for student learning, yet there is little empirical evidence to suggest that this is the case. For many years typical physics laboratories have included confirmation style experiments: experiments which the students carry out to confirm a physical law or witness some phenomena. Labs like these tend to hold the students’ hands through the procedure, causing eyes to gloss over and encouraging students to race through the material as quick as possible to get their points and get out. Recent research has suggested that this style of lab is not particularly helpful to the student, and suggestions have been made that promise to transform the lab in such a way to promote student engagement with the material, and prompt critical thinking.

Inquiry labs, as reformed labs are sometimes called, aim to remove the amount of handholding faced by the student, and provide an experience that is more reminiscent of actual science. Students are given open ended questions and asked to justify their responses based on their findings. Only necessary background material is given, and a handful of suggested tasks to get the student started. For assessment, students can be asked to demonstrate their knowledge by performing a novel yet related task known as a performance assessment. In this experiment, a lab manual which employed confirmation experiments was updated to utilize inquiry based labs with the goal of measuring differences in student performance between the two manuals. The new labs gave less material to the students initially, and prompted them to either research or encounter ideas on their own. In this way, it is thought that students would have a more active role in constructing their learning; becoming more proficient in the subject area and exhibiting improved critical thinking and reasoning abilities.

Measurement of these effects was attempted using two types of assessment: post-lab quizzes to gauge conceptual understanding, and a research based reasoning test given at the beginning and end of each semester to gauge critical thinking and reasoning ability. Quizzes were written to be at roughly the advanced high school level for the subject in question, and were focused on concepts presented or encountered in the lab. The Lawson Classroom test of Scientific Reasoning was used to measure reasoning, and is a multiple-choice test which requires students to select the correct answer to a question and the correct justification to that answer in the following question to get points.

The major goal of this experiment was to measure the effect that changes to a lab manual had on the students taking the lab course; looking for differences between confirmation style labs and inquiry style labs. This required that a confirmation style lab first be measured, and then updated to an inquiry based lab. Previous research has done little to justify the existence of laboratory activities, so this
experiment seeks to add to that body of research, either in support or in detraction of the laboratory activity.

In the following pages, a literature review of relevant research is provided to give the reader a background in previous physics education research (PER) surrounding teaching laboratories. Following that is a description of the methods of the experiment, providing details about the changes to the lab manual and course structure, the different assessments used, and the analysis of the resulting data. Results are then presented along with a discussion of the data and implications. All of this is drawn together at the end with a section on limitations of the study, and avenues for further research.
LITERATURE REVIEW

For as long as people have been attempting to understand physics, people have been attempting to pass their own understanding on to others. This transference of knowledge can take on many forms, such as lecturing, thought-experiments, proposed problems, demonstrations, and guided physical experiments. As time has progressed, we as a people have found new, innovative, more efficient ways of understanding the physical world. It can be argued that we now know more about the vast realm of space and the extraordinarily small domain of the atom than we know about our own oceans and deep jungles. The methods and practices used to conduct science have undergone entire paradigm shifts and subtle refinements in order to get to this point, so it begs the question: why has our method of teaching not undergone similar changes?

It could be the case that scientists figured out all there was to teaching back when we had glass spheres holding the stars in the sky, but this seems unlikely. Only now is the same attention paid to teaching that has been paid to the physical world all these years. There are a number of reasons why: the inhomogeneous nature of the human mind, the hardship of meaningfully quantifying human speech, the lack of willing subjects and researchers, and the perplexing responses given by introductory physics students. As more attention is directed toward these issues, solutions begin to arise. There are now assessments which purport the ability to measure a student’s understanding of conceptual physics through a series of carefully engineered questions, surveys which can assess the expectations and attitudes of students toward physics and science in general, new ideas that seek to outline the best practices for creating curricula and lecture notes, and teams of researchers interviewing confused students to isolate where misconceptions and misunderstandings arise when learning physics. All of these efforts seek to improve the quality of physics education, so that future students come away with a better understanding and a deeper appreciation for physics.

Even with these advancements, much of the research is slow to be adopted by instructors. There are yet areas in physics education which have not received the scientific scrutiny that other areas enjoy. One area that is neglected in PER is the area of the laboratory activity. Many institutions have laboratory classes; often they are offered to students outside of the physics department in order to fulfill a requirement. The idea that a student should have at least some experience with measuring physical variables and calculating important physical results makes intuitive sense to anyone who has worked through the physics curriculum before. After all, laboratory was a core part of how they were taught physics. The perpetuation of the laboratory has taken place despite most lab equipment being costly to store and maintain, and the course itself requiring a significant amount of time from faculty to teach and keep current. It is unfortunate, then, that the little research devoted to understanding physics laboratory has only shown that it is ineffective at developing an understanding of physics in the student. Why then are physics laboratory classes so prevalent in physics education?
The aim of this section is to provide a few sources outlining the current state of physics laboratory, the goals of the physics laboratory, and the possible future of physics laboratory. The American Association of Physics Teachers (AAPT) defines 5 goals for introductory physics laboratories: (to teach) the art of experimentation, experimental and analytical skills, conceptual learning, understanding the basis of knowledge in physics, and developing collaborative learning skills (AAPT, 2014). It should be easy for an educator to agree that these are sound goals; but it is often difficult to evaluate progress toward these goals. To measure progress toward a goal, some type of assessment must take place, commonly something like a lab report or an examination. Depending on the amount of time taken to develop the assessment tool, an educator can have varying degrees of confidence in the inferences made using the results of the assessment given. When educators talk about the validity of a test, they often do not realize that they are not talking about the test itself, but rather the conclusions that can be drawn from the results of that test (Popham, 2000). Thus it is important to consider when making an assessment to consider what it is that you want to assess, and that you design the assessment to meet these criteria.

In physics education there are a litany of different assessments available (AAPT, 2017), claiming to measure everything from conceptual understanding of mechanics and electromagnetic theory (FCI (Hestenes, Wells, & Swackhamer, 1992), BEMA (Ding, Chabay, Sherwood, & Beichner, 2006)), to tests of scientific reasoning and laboratory skills (Lawson (Lawson, 1978), MILI (Galloway & Bretz, 2015)). The Lawson Classroom Test of Scientific Reasoning is generalizable to most laboratory classes, while the MILI is focused on chemistry laboratories, making it less useful in physics laboratories. Many of these instruments are research based, and are currently the best tools available to measure the understanding of our students and make comparisons between groups. It is certainly possible to measure the understanding of students using a home-built assessment, but one must be very circumspect about the inferences drawn when presenting results to others. It is helpful, for the purposes of comparison, to use instruments that have been developed with a research intent in mind, even if there are some drawbacks (Pratt & Hacker, 1984).

Once measurements have been made, how far do students come in reaching the goals presented by the AAPT? On the average, not very far (Lunetta, Hofestein, & Clough, 2006) (Buck, Bretz, & Towns, 2008) (Arons, 1993). Laboratory activities that are confirmation style, or “cookbook”, in nature could be the culprit. These laboratories do not allow the student to develop or execute experiments themselves; they simply provide a recipe to be followed. As put by Arnold B. Arons, "lab could be said to teach only ‘inert ideas’, those which are merely received into the mind without being utilized, or tested, or thrown into fresh combinations” (Arons, 1993). This type of laboratory does not advance the goals set forth by the AAPT, only the goals set forth by the university offering the lab. A synopsis of laboratory education research noted that "... often students do not have clear ideas about the general or specific purposes for their work in science laboratory activities… [they] perceive the principal purpose is following the instructions or getting the right answer." (Hofstein & Lunetta, 2004)
physics lecture is beginning to see improvement motivated by research in defining, communicating, and meeting the goals of the course, the physics laboratory is only beginning to catch up. A group of researchers who developed a rubric to determine the amount of research driven modifications present in a lab says; "... while many new manuals incorporate novel concepts, new techniques, and different equipment, there was no innovation in incorporating inquiry." (Buck, Bretz, & Towns, 2008) What then is “inquiry”, and why does it matter if it incorporated into lab?

Inquiry has many definitions in scientific literature, so many that there is no real concrete, usable definition that should be presented here (Buck, Bretz, & Towns, 2008). Instead, inquiry will be used as a descriptive word in this paper which defines a lab as having more open-ended exploration and less scaffolding, with a focus on critical thinking. As a point of contrast, many current lab manuals can be used as examples as the opposite of inquiry labs: labs which have the students work through a series of well-defined questions which the instructions for solving are clearly defined for the student. There is little thought required of the student; they simply have to turn the crank and write down the resulting number in order to satisfy a requirement placed on them by the university (Hofstein & Lunetta, 2004). This type of lab, the inquiry-less lab, has been shown to meet few, if any, of the goals presented by the AAPT, while preliminary attempts at inquiry based lab show promise in meeting these goals (Barak, Ben-Chaim, & Zoller, 2007) (Cox & Junkin III, 2002).

What makes the inquiry approach different? The goal of most inquiry labs is to encourage students to approach the lab in a more scientific mindset; to pretend as if they were actually doing scientific research instead of just sliding blocks down ramps. Pioneer manuals such as the ISLE (Etkina & Van Heuvelen, 2007) and PSL (Frodermann, 2012) manuals have students design their own experiments to answer a question posed to them. They are asked to justify their work, even if it does not match their expectations. The goal of this type of lab is clearly to reduce the amount of scaffolding, while increasing the amount of time students spend cognitively and physically engaged with the material at hand. This type of learning falls under a key category of the constructivist model of meaningful learning, which posits that in order for learning to take place the student must be engaged in the cognitive, affective, and psychomotor domains (Novak & Gowin, 1984). Inquiry labs attempt to engage all of these domains by making students think critically about the material, giving them more control over their laboratory experience, and letting them work with physical objects to construct their own understanding.

It seems then that the art of experimentation, experimental and analytical skills, understanding the basis of physics, and collaborative learning skills are taught by inquiry labs in a manner that they were not in cookbook labs. Do these labs show any improvement in conceptual understanding over cookbook labs? One problem that lab class research has faced is the lack of assessments designed for laboratory classes (Galloway & Bretz, 2015) (Lunetta, Hofestein, & Clough, 2006) (Lunetta, Hofestein, & Clough, 2006). Since laboratory classes tend to be paired with a lecture, new content often isn’t presented in the laboratory. There are also limitations on the range of material that can be covered in
a laboratory class; lectures usually take place 2 or 3 times a week, homework is assigned, there is a book to read, while laboratory classes usually only meet once a week, and all of the material covered has to be based on the specific experiment that will be performed that week. When an educator attempts to pack more and more material into a 2-3 hour lab class, the depth of knowledge achievable in that timeframe is reduced. So, while many topics could be covered topically in a lab class, it is unlikely that any deeper understanding, the kind of understanding that many physics assessments test for, can be imparted with a laboratory activity. How then do we measure the conceptual understanding of students that is derived from the laboratory class?

A possible solution would be measuring student ability in other areas, and then making inferences about the conceptual gain garnered by students. A similar study measured student conceptual understanding of core physics concepts, namely forces and electricity & magnetism, alongside their scientific reasoning abilities. All assessments used were well known instruments in the PER field, so comparisons to other studies could be easily made. The researchers found that while two different groups of students had very different conceptual understandings of physics (namely due to the manner in which they were taught), their scientific reasoning abilities were roughly equivalent. It would seem, then, that conceptual knowledge is not a good indicator of scientific reasoning (Bao, et al., 2009). It could be the case that higher levels of scientific reasoning relate to higher levels of conceptual understanding, but this link has not been well researched as of yet. One reason for this is the difficulty of measuring and quantifying “scientific reasoning”.

Scientific reasoning could be used as a catch all term to consider all types of constructs which fall under the purview of science and reasoning. Common displays of scientific reasoning abilities are overt skills like estimation, proportional reasoning, and the ability to develop testable questions that would provide insight into a phenomenon. While these skills can be tested, it would be very helpful if researchers could measure the covert skills which are responsible for scientific reasoning, the most relevant of which could be described as metacognition. Broadly speaking, metacognition is thinking about thinking, or becoming aware of one’s own thoughts. Reflection on one’s thoughts is a very important skill in physics, as it is useful for solving problems and evaluating if a result is reasonable. Since asking about metacognition usually alters the metacognitive state, much like measuring the spin of a particle, it is difficult to accurately determine the level of metacognitive activity in the wild. Many studies have attempted to quantify metacognition in students by looking for metacognitive transitions, and researchers have stressed both the importance and difficulty of such a study (Kung & Linder, 2007). Critical thinking is a skill that is derivative of metacognition, and slightly more overt, so it can be used as a proxy for metacognition with some success (Barak, Ben-Chaim, & Zoller, 2007) (Kung & Linder, 2007).

As mentioned previously, inquiry labs are designed in such a way that they create a learning environment which encourages critical thinking. In addition to creating this learning environment, lab manuals can be worded in such a way as to promote “metacognitive transitions” in students. A
metacognitive transition happens whenever a learner changes their behavior in light of a metacognitive thought. An example of this would be a student arriving at an answer for a problem, say finding that the surface gravity of earth is $98 \, \text{m/s}^2$, noticing that the answer does not look correct, and reviewing their work for errors. A student without this metacognitive wherewithal would write down $98 \, \text{m/s}^2$ and just assume there was something wrong with gravity that day. In an attempt to promote these metacognitive transitions, lab manuals can be written with language which prompts reflection (Kung & Linder, 2007). As students progress through the class, one could hope to see self-reflection occurring more frequently even in the absence of prompting. While a promising idea, a group tried to produce metacognitive transitions through the wording of the lab manual did not find any consistent way to do so (Kung & Linder, 2007). It is still noted as a worthwhile endeavor to try and promote metacognitive activity in students in order to improve reasoning and critical thinking skills.

In addition to reflection while working on lab activities, studies have shown the benefits of self-reflection after the lab activity. Not only do students show improved critical thinking abilities, but conceptual understanding also appears to increase after students are given reflective assignments following their usual coursework (Gandhi, Zaniewski, Reinholz, & Dounas-Frazer, 2016). If combined with inquiry based labs, reflection would likely lead to further improvements of a student’s ability to reason and enhance their understanding of the material.

An inquiry based lab with an emphasis on collaborative work coupled with reflective assignments would appear to cover all of the goals presented by the AAPT. This new lab schema would be a non-trivial departure from the current state of lab, but there are resources and guidelines that would help ease the transition (AAPT, 2014) (Etkina & Van Heuvelen, 2007) (AAPT, 2017). Any interested educator should keep a few things in mind that pop up wherever these reform efforts are concerned: goals for the change need to be clearly defined, a guiding coalition needs to be formed to lead the change (the more high level faculty and administrators the better), resources that deal with the implementation and maintenance of such a change should be available, and the proper space should be available (Kotter, 1995) (Knaub, Foote, Henderson, Dancy, & Beichner, 2016).

To address each point in order, goals are exceedingly important in order to have effective change. Not only do goals set the course for change, they provide milestones and small victories which keep the change effort going. As mentioned above, the AAPT has a list of goals for the undergraduate laboratory class, and it would be reasonable to adjust any lab with these goals in mind. Once goals have been chosen, plans should be made which aim to meet the set goals, and a group should be put together to pursue those goals (Kotter, 1995). Change is not the focus of this literature review, but suffice it to say that a motivated group of individuals with support from the rest of the faculty and administration is key to creating and sustaining changes.

Lab space and resource availability can also be touched on briefly. The space in which a laboratory class takes place is very important for shaping the learning environment. Naturally groups should be
seated in such a way that all members can interact and contribute. Studies have shown that 3 is an ideal number of students to have in one group, and groups seem to operate well when placed in groups of three, for a total of nine students at a table (Beichner, 2008). There are different theories as to what students should be doing in the groups, such as assigning roles for the students to follow, making sure that computer time is shared, or switching up groups every so often, all things that could be considered when designing the space in which the lab is to take place. It is also important to make sure that every student has proper access to the lab materials for the day, and that these materials are understandable by the student without much fuss (Gandhi, Zaniewski, Reinholz, & Dounas-Frazer, 2016).

Resources for labs are becoming more and more common: the market for laboratory equipment is growing into prepackaged labs, computer aided data collection, and simulations. There is also plenty of open source content on the internet that can be sifted for gems. Groups producing manuals such as the ISLE labs previously mentioned, as well as rubrics for designing lab courses with inquiry in mind have their materials available to fellow educators (Buck, Bretz, & Towns, 2008). A few noteworthy things that could be considered self-evident but nonetheless are mentioned in research: labs tasks should be calibrated to the caliber of student experiencing them, it is easier to start with scaffolding and then gradually remove it than to start with no scaffolding, labs should be kept relatively simple so that students are not fumbling around with too many concepts at once, and educators should be sure to be active in student learning through the use of discussion tools like Socratic dialogue (Lunetta, Hofestein, & Clough, 2006) (Arons, 1993). Incorporating all of these ideas into a single lab class can be challenging, especially given the time constraint faced by most lab classes, but it is suggested that through these ideas lab classes can become more worthwhile.

Despite the progress being made, there are still a number of unanswered questions concerning laboratory activities. There have been difficulties in justifying the costs of labs by measuring the impact they have on conceptual understanding of physics. One possible reason for this is the lack of assessments designed to be used in a laboratory setting, as many labs do not seek to teach new material themselves. There is also work to be done quantifying the impacts that inquiry based labs have on all manner of variables, namely conceptual understanding and critical thinking ability. There is also plenty of work to be done in misconception research, quantifying the effect of demonstrations on understanding, and comparing virtual labs to corporeal labs. There have been many calls for research in the realm of laboratory physics, with marginal progress made between 1980 and 2002. The void is starting to fill but there is yet work to be done.

The primary goal of this experiment is to add to the body of data concerning laboratory classes. Recent experiments have shown no link between traditional laboratories and improved conceptual understanding, however reformed laboratories were not investigated, nor were scientific reasoning capabilities tested (Holmes, Olsen, Thomas, & Wieman, 2017). Preliminary data gathered from this experiment could be used to guide further research, and will provide insight on how to improve such an experiment to better answer the question: “Are physics laboratories worth it?”
METHODS

To create a control-experiment model, the fall semester of 2016 used a lab manual nearly identical to the one used in spring 2016, except that it was digitized and updated to correct some errors found in the previous year of teaching. Content was still centered around Excel, and the assignments remained very similar to their previous iteration. Changes were made to the way group work was assigned and collected; now groups of students would all work on the same document and turn in one document per group, instead of turning in one lab assignment per student. Post-lab quizzes were introduced to provide a second source of individual measurement in the lab for grading purposes, and as a data source for the research project. This new lab framework would form the basis for changes in the spring 2017 semester, where changes would be made to the lab manual’s content instead of just the style.

Data Collection Design

Data was collected in three different areas: conceptual knowledge, reasoning ability, and demographic information. Conceptual knowledge was tested at the end of each lab in the form of the post-lab quiz. Each quiz consisted of 10, 5 item multiple choice questions. Of the 13 quizzes administered throughout the semester, 9 had items which were used for this research. A complete list of content questions asked can be seen in Appendix A. A more thorough explanation of the development and administration of these post-lab quizzes is given in a later part of this chapter.

Reasoning ability is difficult to test, as has been discussed in the literature review. The Lawson Classroom Test of Scientific Reasoning (LCT) was selected due to its previous use as an instrument to measure reasoning. This test contains 24 questions, 22 of which occur in pairs and 2 that are individual items. The test is scored out of 13 points, 1 point given for getting both items in a pair correct, and the other two given for the last two questions. Various aspects of scientific reasoning are measured by this instrument, and the paired questions require students to first select the correct answer to the stem of the question and then chose the correct justification for the answer. This design reduces the chance that students will correctly guess an answer, and can tease out inconsistent ideas in student thinking. Ceiling effects have been observed in the assessment before (it is not terribly difficult), but it is the current standard for this type of measurement. While Chinese translations of this test do exist, it was decided to administer only the English version, as it is difficult to ensure that the translation functions in the same way that the original does. There were concerns that the length and complexity of the wording of the test could have differential effects on some populations, but there was not much to be done. Unlike the post-labs, which were assignments that the students were required complete for points, the LCT was given at the beginning of the first and last class as an extra-credit opportunity. The idea here was to encourage the students to do their best, and avoid penalizing
them if they did not excel. The Lawson Classroom Test of Scientific Reasoning is a controlled document, and as such has not been included in the appendices of this thesis. Those wishing to see a copy of the test can find it online at a number of different university websites, or on Physport.

Demographic information was collected for each student, and the questions asked can be found in Appendix B. Information was collected on the student’s year in school, gender identity, ethnicity, whether they consider themselves international, previous physics/laboratory experience, and motivation for taking the course. In addition to this information, the section and time of day which the student took the course were also recorded. All of this information was kept to show that populations of both semesters were comparable.

Post-Lab Addition

Post-labs were designed to be consistent between semesters, and were given as 10, 5-item multiple choice questions. Broadly speaking, these post-labs were written at the advanced high school level, as the course is for students who have not had a physics class at the college level before. Old SAT physics questions and other online study tools were used as inspiration when making the post-labs, and the learning outcomes presented by the Ohio Department of Education (Ohio Department of Education, 2011) were kept in mind when constructing assessments. Questions were written while being mindful of the best practices laid out by Popham in his book, Classroom Assessment: What Teachers Need to Know (Popham, 2000), to improve the overall efficacy of the item for both assessment and teaching. Examples of best practices include shuffling the location of the correct response answer, having uniform response lengths, having short response options, avoiding hints to a question in other questions, using understandable language for the level of the student, and minimizing bias in writing as much as possible (i.e. not using culture specific examples). Each post-lab was designed in such a way to have 2-3 questions per core concept of the lab. Attempts were made to have varying levels of difficulty for each concept, and throughout the quiz as well. Earlier iterations of the post-lab had more observational questions (i.e. which of the following measuring devices was NOT used in today’s lab?), and these were largely phased out and replaced with content specific questions as more quiz making experience was acquired.

Quizzes were delivered using Canvas infrastructure, which has pros and cons. Canvas allows students to take the quiz on their own time, in their own home; but it also allows students to take the quiz on their own time, in their own home. Initially, quiz access was limited to 15 minutes, but this was increased to 30 minutes halfway through the first semester in response to student requests asking for more time to be able to read the questions. Students had slightly more than 24 hours after completing the lab activity to finish the quiz. This was done so students had the relevant material fresh in their mind when they went to complete the quiz. This was kept consistent between the two sections in the second semester, but varied in the first semester due to variances in teaching style.
Content Modifications

A series of changes were implemented over the course of the spring semester which were intended to reduce the amount of confirmation in the lab, and increase the amount of student engagement with concepts and novel material in order to promote critical thinking and conceptual understanding. Labs, in general, were designed to include no more than 3 main content topics, often fewer, to reduce the amount of non-essential information to which students were exposed. In labs which had material students were familiar with extra information did not pose much of a problem, but in later labs where students were working with things like physical optics and circuits, too much information can cause students to become overwhelmed and disengaged. Attempts were made to reduce this effect by parsing the amount of material the students are presented in the introduction, and reducing the number of concepts and ancillary encountered in the later lab activities.

Care was taken to write the introductions in plain English, and they were designed to be consistent between topics. Each lab introduction contained material in the same order, and in approximately the same volume for each lab. Three key features were added to each introduction: rough learning outcomes were added at the beginning of each lab, questions designed to direct student thinking in helpful ways were added to the body of the introduction, and summary points were added at the bottom of the introduction which encompassed the main themes to consider when working on assignments. Overall accessibility was improved, breaking up chunks of text and organizing content under consistent categories. Links were added to take students to the correct assignment documents and submissions. Diagram, picture, or other media use was kept as similar as possible to the fall semester.

Pre-lab assignments were made unique for each lab, instead of being a generic form that was used in the fall. The first and last questions were kept the same (what is the object/ask a question), but the middle section of define terms/measurements/calculation was replaced with content specific questions. Comparisons of these two forms are given in Appendix C. Some questions prompted students to do extra research on their own, as the answer was not stated directly in the write-up (but could be inferred, or found with a quick google search of the appropriate phrase). Each question attempted to focus students on the concepts which would be present in the lab and in the post-lab. These questions were rarely demanding, and could be completed in less than 5 minutes after reading the introduction. Pre-lab questions were easier than post-lab questions, except in the case of simple post-lab questions.

The in-class lab assignments were reworked to include a smaller number of concepts, keeping with the theme of the introduction and the post-labs. ISLE labs (Etkina & Van Heuveln, 2007) were used as reference material whenever topics in the class overlapped with extant ISLE material. Questions which were simply rote calculations or recitations were removed, in favor of questions which required more effort from the student. Previously this course had an emphasis on scientific Excel, which was
a major source of disengagement for many students. As the labs went on, students would collect the requisite data and then leave the excel work to the most practiced group member. To remedy this situation and promote group work, Excel is absent from the early lab assignments in this course. Only once students have had experience with the laboratory and thinking about physics are they introduced to scientific Excel. Quantitative skills were thus introduced after more qualitative skills have set root, to reduce competition for student brain space.

As a result of the aforementioned changes, labs often had a smaller number of questions but the average response to each question increased in length. To give student groups the opportunity to make a decision and support their decision with information gathered in the lab, questions were often open ended, and required students to give at minimum two examples or pieces of supporting evidence: commonly students would justify their answer with an explanation, or identify a possible source of error. It was made clear to students that “human error” was an insufficient source of error, and that they must give an explicit example such as “human reaction time become relevant on this time scale” or “parallax error”.

After a number of this type of question, many labs contained some sort of performance assessment, in which students would use the concepts they were working with in earlier parts of the lab to tackle a semi-novel task. Examples of this include creating an equilibrium situation, shining a laser on a penny in a tank of water, or explaining a demonstration of electromagnetism. These tasks are accomplished by the group, and the group is scored not only on their ability to perform the task, but also on their cohesion. These tasks are intended to promote student engagement by providing a task which is more grounded in the “real world” instead of arcane physics land. Critical thinking and communication are also expected to result from this sort of task. To further these goals, student groups were also switched in the middle of the semester to give students experience working with new people.

Appendix D includes a comparison of the fall and spring lab assignments. Questions were kept similar to maintain consistency, though some questions were removed in favor of others. Notice that the blocks of explanatory text in the fall semester were largely removed for the spring semester; much of this information was removed and left for the student to sort out, while the critical pieces remained or were placed in the introduction. Question 6 in the fall includes the introduction of a performance assessment, where students were to calculate the incident angle required to hit a penny submerged in a tub of water. This exercise required that students use simple trigonometry and Snell’s law together, in a way that was not immediately obvious to many of them. For those students who were not successful, there was an opportunity to talk about the reasons why; the effect of the plastic walls which were often omitted from their calculations, or the use of the incorrect angle with the protractor, or the difficulty in setting up the laser to shine along the correct path. The last section on image projection was reworked to remove any references to the lens maker equation, which previously was a source of confusion for students. Instead they were asked to play with the location of the various optical
elements to see what changes could be made to the image. While less scientifically rigorous, this process is more like actual experiment when compared to following explicit instructions.

Section Consistency

For thoroughness, it seems appropriate to note here that there were four sections of this class studied, with two different instructors teaching. It was also anticipated that there could be relevant demographic differences between some of the sections, so demographic information was collected. Differences between instructor style were also anticipated, but no attempt was made to strictly control for them. The classroom is a fluid environment, and attempting to get two different instructors to conform to the same protocol for teaching is an expensive endeavor. Both instructors did teach from the same resources, and the instructor who taught first typically met informally with the instructor who taught second to let them know of any changes made from the last semester, what roadblocks to look out for, and how to best manage any of the new equipment.

Analysis

A variety of different analytical techniques were used to summarize the data, ranging from simple descriptive statistics to a Rasch model. For each post lab a bar graph was created comparing percent correct responses for each item between semesters. In addition to these bar graphs, Rasch analysis was used to gauge the reliability of the post-labs by investigating item ordering and the Cronbach Alpha value. Two-tailed t-tests were conducted on the post-lab data as well as the gain scores on the LCT from the fall and spring semesters to check for significant differences between the two sampled groups.

The Rasch model falls under item-response theory, and attempts to quantify the ability of a participant as well as the difficulty of a test item. The underlying idea is that not all questions are created equal, and neither are all test takers: there should be some consideration in a statistical analysis for this variation. There are a variety of other features of the Rasch model that appeared useful, but in studying this model it was difficult to discern if it was appropriate to apply to the data collected in this experiment. The model requires that only one construct be tested in order to make valid inferences from the output. It could be said that “physics” is a single construct, but there are myriad different aspects to physics; so many so that it no longer seemed appropriate to use a Rasch model on the combined data set. In light of this information Rasch was used to created wright maps, which order the items of an assessment based on their difficulty, for each post-lab. If these items occur in the same order for both fall and spring semesters, it is implied that the quiz was functioning in exactly the same way for both groups, and that the two groups are acting in the same way. If the treatment had no effect, it is expected that items would occur in the same order. The Cronbach Alpha for each post-lab is also generated, and is a measure of internal consistency for an instrument. A measure of .7 or greater
from this test indicates that the items of the assessment work together to measure the same construct(s).

Institutional Review Board Application

Since this is an experiment involving human or animal subjects, a review by an Institutional Review Board (IRB) was necessary to begin this project. For simplicity, this study only included those students who were over 18 years of age and gave their consent to participate in the experiment. Since the material used in the experiment was all part of the normal operation of the course, and any identifying information collected from participants would have their name stripped with little risk of harm to the participant, a level 1 application was filled out. This type of application is an expedited application which assumes low risk, non-intrusive experiment design. Dr. Jennifer Blue was the primary investigator on the application.

Required for this application were a variety of supporting materials, including a consent form and a statement of purpose. The experiment was presented as a study on the effectiveness of changes made to a laboratory classroom. Consent would be gathered from students at the end of the semester via a consent form, a copy of which is included in Appendix E along with the IRB application. Students were given the consent forms at the end of the semester, along with a letter explaining the research and the LCT with a demographic inventory. It was made clear that participation in the research study was voluntary, and that there were no penalties for deciding not to consent to have their data used in the project.
**Data**

There were 101 participants in the fall, and 94 participants in the spring. Not all of these participants completed every assessment, so the number of involved individuals is given with each figure. Demographic data is given first, followed by data from the post-labs, and lastly the data from the LCT. Those interested in seeing similar figures many times are encouraged to read the paragraph under each heading for a flavor of the data, and then proceed to the discussion chapter, where all interesting features are summarized.

**Demographics**

Included on the following pages are the results from the demographic surveys from the fall and spring semesters. Data is presented per semester, not per section, as differences at the sections level are covered in the discussion section when relevant. Physics 103 is characterized by the high percentage of international students, many of whom register for the class as it is still open when they start registering. Most students are in their first or second year, though all class levels are represented. Many students come in with some basic background in physics and in a laboratory setting. There is no prevailing reason for which all students take the course, and males and females are nearly equally represented. The following two pages contain graphs of the demographic breakdown by semester.

Figures 1 and 2 on the following pages illustrate that the demographics of each semester are similar enough that it is not unreasonable to compare them. Similar sample sizes and similar population characteristics provide for summary statistics to be presented for the semester as a whole, instead of by section or by demographic. A slight difference in the number of physics courses and lab courses taken by the students in the fall semester is present, but this difference did not appear to influence the results of the LCT pre-test, as will be discussed later in LCT section.
Figure 1: Fall semester demographic data, N=96
Figure 2: Spring semester demographic data, N = 94
Post-labs

In looking at the question breakdown for the post-labs, the general trend seems to be that the students in the spring semester performed slightly better on the earlier labs than did the students in the fall, and slightly worse on the later labs. Two of the post-labs showed significant gains between the fall and spring semesters that were likely caused by reformed lab pedagogy. A few questions have large discrepancies between fall and spring semesters, though there does not appear to be a common thread connecting them all; they are not all purely conceptual, nor do they all relate to lab content that was rewritten. A clear trend does not emerge where one group did significantly better than another group, though there do appear to be certain labs in which one group outperformed the other. Large discrepancies are addressed as encountered.

Rasch analysis of the post-labs was not particularly inspiring. None of the quizzes had a Cronbach-Alpha value greater than the standard .7 for internal consistency. Nearly all of the quizzes also had a standard error of measurement greater than 1, which is especially poor. There are a variety of limitations to consider: the program was only able to handle 75 participants, so 75 participants were chosen at random to meet this limit. While not an ideal method, Rasch analysis is not overly sensitive to outliers, and a sample of 75 people should be enough to provide a rough estimate of a population of less than 100. The small number of items for each quiz limits the internal consistency, as there are fewer points to cross reference. Many items were also quite easy or quite difficult, which makes it difficult to stratify the questions into different categories; they tend to lump instead of forming tiers. It is a small consolation that in many cases the Cronbach-Alpha scores were similar between semesters, so it can be tenuously claimed that the instruments functioned similarly for both groups.

The wright maps produced by the Rasch analysis paint a different picture. Typically, when you give the same instrument to two similar groups of students, you would find that the order of difficulty of the items on the instrument is roughly the same. The actual difficulty may vary, one group may find question 2 to be easier than another group, but both groups would find that question 2 is less difficult than question 4, but more difficult than question 7. Comparing the results between two groups is a form of test-retest reliability, and can be useful when you want to administer your test to other groups and compare them to your previous groups. Interestingly, the wright maps produced by the analysis of the data collected very rarely have the items in the same order. In some cases, the item ordering is the same when you take into consideration the error margins, but in others the items are simply in no comparable order. This could be interpreted in a number of ways, the main two being: issues with test design, number of questions, or implementation, or the two groups are dissimilar enough that the items had different difficulties for each group. Given the poor, but consistent Cronbach-Alpha measures, it could be the case that each semester had students which performed with different strengths, as influenced by the changes to the lab manual.
Reading a Wright Map

Interpreting a wright map is not all that difficult. From top to bottom, items and persons are listed in order of descending difficulty or ability. A person toward the top of a wright map did very well on the assessment, and an item toward the top of a wright map was very difficult for most persons. If a person and item occur at the same spot on the scale, this indicates that the person would get that item correct 50% of the time. The further the person is above an item, the more likely they are to get that item correct. These are logarithmic probabilities, so the increase in percentage is not linear.

The left side of the map lists the persons, or individual data points, used in the analysis. Typically, these persons have a roughly normal distribution, depending on how difficult the items are. Those persons who are higher on the scale are more competent at whatever the topic of the assessment is than those at the lower end of the scale. Items are presented in a similar fashion, and are ideally evenly spaced so that there are a variety of difficulties present in the assessment. Wright maps are presented here to compare the location of the items in the spring and the fall. If both groups were experiencing the assessment in the same way, we would expect to see the items occur in the same order, preserving their difficulty between semesters. If groups experienced items differently, the locations of the items would change between semesters. In figure 3 we see that the ordering of the items is not the same: Q2 = Q3 > Q5 > Q4 > Q1 in the fall, compared to Q5 > Q2 > Q1 > Q4 > Q3. Questions 5 and 1 were more difficult for students in the spring compared to the fall, while question 3 became easier.

Many of these maps report a large standard error, which is in part due to the small number of items and the fact that they were coded in binary form (0 = incorrect, 1 = correct) instead of their raw form (with 4 incorrect options and 1 correct option). This large error should be kept in mind when interpreting these maps, as some items which appear to have shifted by a large amount in the spring are still within ~.5 of their order in the fall. While not compelling on their own, wright maps such as these offer additional perspective on the function of the conceptual assessments.
Figure 3 contains the Wright maps for the Measurement and Uncertainty post-lab. Looking at the right half of each map, the question items do not occur in the same order, indicating that the groups found the questions to be of different difficulties. One possible explanation for this could be the transition from a numeric answer to multiple choice from the fall to spring, which would lower the difficulty of questions 1, 2, and 3; this seems consistent with the maps. Question 5 was especially difficult in the spring; this is likely due to the removal of the word “all” which may have served as a tell in the fall.
Since the quiz specifics underwent significant changes from the fall to the spring, it is not unexpected that the α value decreased considerably; many students did significantly better on this quiz as seen in the higher position of many persons on the left half of each map, and as seen above in figure 4. This may have led to a ceiling effect, reducing the internal consistency of the assessment. It is likely that the changes between semesters in figures 3 and 4 are to be attributed to the change in assessment instead of the change in lab material.

Figure 4 also illustrates the difference in student success in the spring and fall. The addition of multiple choice questions made Q1, Q2, and Q3, easier in the spring, and the removal of a telling phrase in Q5 made it more difficult. Changes in wording may also be the cause for the improved correct response rate in the spring for Q4, or a greater portion may have been familiar with mercury thermometers and mercury. A t-test comparing the two semesters suggests that there is a significant difference between the groups, with a p-value of .003. In general, a p-value of less than .05 suggests significance, indicating that there is a less than 5% chance that the two compared groups have the same mean characteristics and are of the same population. In this case the difference is more easily explained by the change in assessment than by changes in material and thus the students.
Standing Waves

```
<table>
<thead>
<tr>
<th>Measure</th>
<th>Person</th>
<th>Map</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>#</td>
<td>T</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>+</td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>0</td>
<td>#</td>
<td>M</td>
<td>5</td>
</tr>
<tr>
<td>-1</td>
<td>+</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>-2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

```

Fall: $\alpha = .3$, SEM = 1.11
Spring: $\alpha = .44$, SEM = .97

*Figure 5*: Standing Waves wright maps

Interestingly, the above wright maps are some of the better formed of the bunch. The low $\alpha$ value is likely due to ceiling effects, as nearly all the students were found to have abilities much greater than the most difficult item, question seven, in both cases. Persons also follow a roughly normal
distribution, which is ideal. The question ordering is much more consistent when compared to figure 3, and much of the disordering can be explained by changes made to the manual. Question 2 was likely easier in the spring as students were given a task which had them create a standing wave on a string with their hands, possibly reinforcing the idea of what defines a transverse wave. Question 1 was likely more difficult in the spring as students were not given a frequency to start with, nor were they tasked with calculating the fundamental frequency of their standing wave generator; they had to determine the correct frequency experimentally. Question 5 was easier in the spring as students were tasked with finding the period of their handcrafted standing wave, perhaps reinforcing the definition and meaning of period in a physics sense.

Figure 6 above illustrates a convincing improvement from the fall to spring semesters. A t-test of the data suggests that there is a significant difference between the two semesters (p = .0178), and it is likely caused by the updated material. All concepts appeared to hold better in students minds as they took the post-lab assessment. While there is a convincing improvement, it is important to note that it is possible that the improvement arose from a lab more tailored to meeting the goals of the post-lab. It is difficult to differentiate for certain between the lab benefiting from general improvements related to the implementation of the material and style of the questions, or if there was a more focused attempt at teaching toward the post-lab.
Simple Harmonic Motion

Results from the simple harmonic motion lab indicate that students performed well in both semesters. Item disordering is less pronounced except in the case of Q5, all others are in roughly the same order when considering the size of the error. Question 5 is a seemingly innocuous question about a floating leaf where students are asked to determine the frequency when given the number of cycles per second.

Figure 7: Simple Harmonic Motion wright maps
It is unclear why this item was found to be more difficult in the spring, though it is concerning as the relationship between period and frequency is a fundamental part of this lab. This is especially confusing when contrasted with performance on question 7, which asks students to calculate the period of an oscillator with a known frequency. It would appear that students were more proficient with material they worked with directly and were asked questions about in the lab, but did poorer at a task which required them to apply the same concepts to unfamiliar situations.

Questions 4, 6, and 7 all appeared to be considerably easier for students in the spring. Instead of a lab which told students the relationship between mass, length, gravity, and period of a pendulum, students were tasked with sorting out the relationship themselves. This seems to have had a large effect in solidifying the relationship between mass and period, as seen in questions 4 and 6. Another plausible explanation is that students would spend more time trying to sort out the relationships before asking a student for confirmation, allowing the ideas to develop further and become more ingrained than they would have if they were just told the relevant information. A t-test suggests that there is a difference between the two groups (p = .037), so some benefit may have been derived from the changes made.

Figure 8: Simple Harmonic Motion post-lab scores
It is unclear why the spring assessment had an abysmal $\alpha$ value, and it is suggested to take any inferences made from that assessment with a grain of salt. A large amount of item disordering is present, and the person ordering is dissimilar as well. Q4 for a rise in difficulty due to a de-emphasis on the particular tab of the simulation which the question referenced, and students in the fall semester may have received some advice to carefully observe that portion of the simulation. Question 1’s rise
in difficulty is disturbing, and may be attributed to the lack of explicit definition of fields in this lab. The previous week, Magnets, explored the idea of a field in depth, and the Generators lab talks at length about magnetic fields, but it seems the concept did not stick well. The question in question, “which is the most similar to a gravitational field,” had an increase in the number of students who incorrectly answered, “Electromagnetic Force” who otherwise did very well on the assessment. This item may have been a contributing factor to the low α value.

While a t-test suggests that a significant difference between these two groups is unlikely (p = .73), the requirements for the test are not precisely met. The large difference in participants between the spring (95) and fall (68) make comparisons difficult. Gains on questions 5 and 8, which address the idea of conservation of energy, are likely due to an increased focus on the topic in both the introduction and the pre-lab assignment. Gains on questions 2 and 3 may be due to the introduction of physical demonstrations of the creation of a current in a coil, but it is difficult to say for sure given the quality of the data.

![Figure 10: Generators post-lab scores](image)
Items from the geometric optics post-lab appear to occur in nearly completely different orders, indicating a significant difference between either the population abilities or the assessment difficulties. Similar $\alpha$ values suggest that the differences lie more in the population abilities than in the item difficulties.

**Figure 11**: Geometric Optics wright maps

Fall: $\alpha = .5$, SEM = 1.44   
Spring: $\alpha = .49$, SEM = 1.39
A t-test returns a p-value of .619, indicating a lack of significant difference of the means, so losses seem to wash out with gains. Losses on question 8 are likely rooted in the rephrasing of the introduction of index of refraction: instead of explicitly stating the phrase as written in the question it was introduced more obliquely, making it more difficult for students to recognize the correct answer. Gains on question 3 are surprising in contrast, indicating students in the spring had a good grasp of the concept of refraction in relation to the refracted angle. Question 4 was likely easier due to increased practice in determining the direction light will bend as it passes through a medium. This lab is one which underwent major changes going from fall to spring, the assignments are given in Appendix D for comparison. With these modifications, it is clear that students performed differently on the post-lab assessment but it is difficult to pinpoint the underlying reasons when their averages are similar.
The extremes on this assessment are interesting. Question 4 seemed to increase in difficulty relative to the rest of the assessment, and question 9 skyrocketed in difficulty. Question 3 saw a drop in difficulty, perhaps due to the rewording of the question. Interference as a buzzword is much easier to identify than to pick out that interference was the main theme of the sound lab which also happens to relate to this lab. Other disordering occurred, but this is not unreasonable considering the error margins.
A t-test of the data suggests there is could be a significant difference at the 5% level (p = .052), but this may be skewed by question 9. Without the large drop in correct responses to question 9, there would not be a significant difference between the two means. Questions 8 and 10 both require students to read a significant amount of text compared to other questions, and then relate the ideas that they just learned to similar but unfamiliar ideas in the question. It appears that the students who did not experience the treatment fared better at this task.
Sporting the best measure of internal consistency of all the post-labs, Atomic Spectra was the lab that changed the least between semesters. A new pre-lab was made, but most of the lab content stayed the same, including lab questions and methods. The upper and lower ends of the wright map are in agreement, with some variation present in the middle. This variation is not as severe as other cases, and does fall within the error margins.

Figure 15: Atomic Spectra Wright maps
The losses are visible in figure 18, with the spring semester doing slightly but consistently worse across the board. A t-test suggests that the means are not significantly different at the 5% level (p=.071). Still, it is curious to note the consistent drops in the spring semester. This is possibly due to sample size, as fewer students were sampled in the spring, but it is not unreasonable to think that those students who did not complete a mandatory assignment may not have done the averages any favors had they completed the assignment. A possible source of this decline is the waning motivation of students to attend and complete lab courses at this point in the semester. The allure of summer had begun to set in, as there were only two class left after this one.
There are some very interesting results to be seen here, namely the locations of question 3 and 6. These questions both ask about the energy content of light; Q3 the most, and Q6 the least. The range of items between Q3 and Q6 in the fall makes one wonder what happened when we see both items on the easy end of the spectrum in the spring. The other items appear in nearly the same order with a few inversions.
Students in the spring did better on 6/10 items, but there is enough give and take between problems for a t-test to report no significant difference (p = .978). This lab was only received minor modifications in terms of questions and methodology, so it could be that the variations present are just the natural ebb and flow of student ability from year to year as no one concept saw net gains or losses. This would also suggest that any benefit derived from nearly a semester of reformed laboratory work did not prepare the students any more for this particular lab.

Figure 18: Photoelectric Effect post-lab scores
Questions 6 and 7 were disordered, likely due to increased emphasis on the relationships present in the ideal gas law as described in the pre-lab and introduction. Otherwise only small differences appear in the Rasch output for this lab.

Fall: $\alpha = .57$, SEM = 1.16  
Spring: $\alpha = .54$, SEM = 1.23

Figure 19: Absolute Zero Wright maps
The two major drops, Q4 and Q9, are likely the cause of poor rewording of the post-labs. Attempts were made to update the post-lab wording for clarity, but in retrospect they became more confusing, a conclusion supported by the data here. This lab saw a large reduction in scaffolding: students were not told the method by which they were to determine absolute zero, instead they were told to think back to the previous experiment where they performed a similar task to determine the work function of a metal. Emphasis on the ideal gas law throughout the experiment improved the correct response rate for Q6 and Q7, but this may have been a result of coaching instead improved lab materials. A t-test of the data for these two groups suggests no significant difference between the averages of the two (p = .446), though differences in certain areas are certainly present.

Looking at the summary of p-values in table 1, there appear to be noticeable differences in 3 of the 9 post-labs between semesters. The difference in the Measurement and Uncertainty lab is almost certainly due to changes made to the post lab than due to the lab itself, so it should not be counted as a lab which clearly improved student understanding. Standing Waves and Simple Harmonic Motion were both early labs as well, meaning their post-labs had more changes than later labs. Even with this in mind, it seems that the changes made to the lab improved student conceptual understanding, but it is difficult to say with certainty that updated pedagogy is the only contributing factor. All of the other labs did not see significant gains or losses on average, though there were cases where certain questions or core concepts saw changes due to the lab manual reform.
Lawson Classroom Test of Scientific Reasoning

Students were given this test at the beginning and end of the semester as an in-class extra credit assignment. Not all students completed both the pre and post test, so only those students who did had their scores included in the analysis of this assessment. The fall semester had 91 participants which met these criteria, and the spring semester had 82. A t-test comparing pretest results of both semester showed that there was no significant difference in the two groups, with a p-value of .71. For reference, a p-value of less than .05 is the general cutoff at which differences become statistically significant, though the t-test should not be the only measure discussed for an instrument. This result supports the claim that both semesters were of comparable reasoning ability at the beginning of the semester.

Comparing the pre-post gain on this assessment for students in the fall semester and students in the spring semester yielded no compelling evidence that there was any significant difference between the two. The average gain on the scored test was quite similar: .75 with a standard deviation of 1.88 in the fall, and .68 with a standard deviation of 1.81 in the spring. This was out of a possible score of 13, so there was an average gain of 5.7% in the fall compared to a gain of 5.2% in the spring. A two-tailed t-test produced a p-value of .81, which suggests no significant difference between the gains of the two groups, indeed the difference is even less than the difference at the pretest.

![Fall LCT Graph](image)

*Figure 21: Fall Lawson Test results*
Figures 21 and 22 show the correct response rate for the spring and fall semesters, pre and post. There is a clear trend in both semesters of gain on nearly all questions, suggesting that the course did slightly improve students’ ability to reason.

<table>
<thead>
<tr>
<th>LCT Results</th>
<th>Fall Pre</th>
<th>Fall post</th>
<th>Spring pre</th>
<th>Spring post</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>7.66</td>
<td>8.41</td>
<td>7.72</td>
<td>8.48</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>2.49</td>
<td>2.34</td>
<td>2.44</td>
<td>2.37</td>
</tr>
</tbody>
</table>

Looking at the averages in table 2, ceiling effects do not appear to be a serious factor in this study: only 5 people scored a perfect 13 throughout all the sections and semesters, 2 in each of the posttests, and 1 on the fall pretest.
DISCUSSION

Results

Previous research has suggested that laboratory courses confer no measurable benefit to student understanding of physics concepts, but that laboratories that are reformed in light of current research may be an exception. This experiment finds a small amount of support for the idea that reformed laboratories are better at teaching physics than traditional laboratories. A handful of topics saw increased conceptual understanding as a result of reformed laboratories, but many other topics saw no significant change. As a general trend, those topics which were related to physical phenomena that students would have been familiar with before the course; forces, motion, and Newtonian ideas, saw more benefit from the reformed lab than did those which dealt with less commonly encountered concepts, like electromagnetism and thermodynamics.

Lab reformation did not have any significant effect on students’ scientific reasoning ability. Both groups came in at comparable levels in terms of reasoning ability, and left with comparable levels of ability. This result suggests that the physics laboratory does improve the reasoning abilities of students. This aspect could be further explored in the future to determine if the gain was from the lab course, or from the previous exposure students had to the assessment, or from the increased experience students gain over the course of a college semester. Measures directly involved showed no difference in student reasoning ability between the two semesters, but anecdotal evidence suggested that the reformed material had a positive effect not only on student reasoning ability but also morale.

Student writing samples were much more cogent in the spring semester when compared to the fall semester, due to higher expectations stated in the questions. Much more attention was paid by students to sources of error and justification for answers; at least as expressed in writing. Performance assessments gave students the opportunity to talk through the physics ideas in a graded environment, allowing instructors to view another side of student ability not previously explored in lab assessments. These effects did not contribute in a measurable way on any of the assessments used in this study, however.

In conclusion, this experiment reveals a small added benefit to conceptual understanding and no benefit to scientific reasoning ability for students experiencing a reformed lab in place of a traditional lab. This experiment was not flawless, and there are number of avenues through which the experiment could be improved in future iterations to increase the gravity of the results.
Limitations

Measures of conceptual understanding were not as internally consistent as would be ideal due to small numbers of items. Since the conceptual assessments were developed just before the fall semester, early implementations required changes in the subsequent fall semester, reducing the amount of material that could be pulled on as data. Items were also kept the same between semesters, and while strides were taken to reduce the occurrence of cheating in the course, academic dishonesty had been a problem in PHY103 in the past. Assessments were also taken online, so it is nearly impossible to know if students were using other materials or collaborating when taking the quizzes. The lack of massive gains between semesters would suggest that cheating was not present on a large scale, but it is still possible that it was present in some individuals.

Many of the problems with the conceptual assessments could be remedied over time with the development of a question bank. A large bank of questions tested over multiple semesters would provide researchers with self-validated tools to measure conceptual understanding. The same question given year after year under the same conditions would begin to crystallize in difficulty score in Rasch analysis. Once a database of previous item difficulties has been developed, Rasch analysis becomes more powerful as one can confidently identify on a wright map when questions are more or less difficult for a new population compared to previous populations, pinpointing changes in the target population. This information would also allow quizzes to be constructed with items of varying difficulty, effectively increasing the range over which the assessment is useful for gathering information. An assessment with 4 questions, one hard, one difficult, one easy, and one simple, would be more useful than a 20-item assessment of easy questions in terms of placing student ability.

Scientific reasoning is a difficult thing to pin down, and as such there are very few research validated instruments available that can produce a reasonable measure. The Lawson Classroom Test of Scientific Reasoning is one such test, but it does not come without flaws. It is quite an old test, first developed in 1978 and only undergoing one major revision since then. In this time it is possible that the test has lost some of its ability to categorize students to the passage of time. The use (or development) of a different reasoning test may provide more reliable results in future iterations of this experiment. Such a test would need to be borrowed from a field other than PER, or developed over the course of many years. Paired multiple choice questions which require the identification of the correct justification to a question are an interesting solution to the problem of measuring reasoning ability, but are difficult to produce as effective distractors need to be researched and implemented.

In developing the new lab material, the old material and post-lab assessments were both kept in mind. It is certainly possible that the new material unconsciously, or in some cases consciously, catered to the material present in the post-labs to improve student performance. Situations where students had complained about the lack of clarity on certain concepts, or the unfair nature of a post-lab question which they felt had not been adequately explained in the introduction may have had unintended effects.
on the new lab material. To control for this, it might be helpful to have a team of researchers working on such a project, where one part of the team works on developing the conceptual assessments, and other works on redesigning the lab manual without specific questions in mind.

While not tested for in the study, there could be interesting differences between demographics and sections. As a whole, there was no significant difference, but it may be that certain groups found benefit in the changes and other groups did not. This is one of multiple avenues open for future works.

Future Works
Investigating the differences between sections is something that is possible with this data, and has been brought to the attention of the Statistical Consulting Center at Miami University. With any luck, the results of that consultation will be included as added to this work. A proposed regression model would identify differences between sections and demographics when looking at the LCT gain and scores on the conceptual inventories. This information could be valuable for making the lab equally effective for all individuals enrolled in the course, or could statistically prove the superiority of the evening instructor over the afternoon instructor.

Collection of student feedback could be valuable for measuring the affective domain. Anecdotally students preferred and were more engaged in the reformed laboratories, and this could be harnessed and developed into gains on conceptual understanding, scientific reasoning ability, or department image. Previous unpublished research by this research group has found a link between student opinion of physics and student ability in physics in a lecture class; it may be the case that this link could be exploited in the laboratory environment to benefit students in lab courses.

Increased emphasis on performance assessment could yield some other result than the one arrived at in this study. Students preferred performance assessments to typical lab work, and this type of assessment may have some benefits to the students if implemented on a large scale. Being resource intense, performance assessments have not been implemented or studied widely, but they are another avenue which could be explored when trying to improve the effectiveness of the physics laboratory.

Including multiple small experiments instead of one large experiment could have a positive effect on student retention of material. Finishing one concept and then advancing to another may be a lighter cognitive load on students than maintaining three concepts at once. Faster paced work could potentially keep students engaged more, and prevent the course from dragging on. An experiment to test this could take a traditional lab and reform it using experiments designed for use in studio physics classrooms, and as such serve as a testing ground for both studio physics and standalone physics laboratory classes
Experiments into the role of the group could also prove enlightening. Though suggestions exist on what the proper group size should be, there is a lack of concrete evidence in this area. A lab manual which taught students effective group work would be a prized gem in any department, and would equip students with useful skills for the future, making them valuable alumni. Little emphasis is currently put on the explicit teaching of group work, and it may be the case that students who are taught how to work together develop a better understanding of the material they are presented with as a group.
Bibliography

https://www.aapt.org/Resources/policy/goaloflabs.cfm

AAPT. (2017). *Inventory of research based physics assessments*. Retrieved from Physport:
https://www.physport.org/assessments/


Measurement and Uncertainty

Q1 (Fall):
Imagine you are a person that is 75kg, and takes up .09m³ of space. Now imagine you are on Jupiter where the surface gravity is 24.79 m/s².
What is your weight on Jupiter?
(Enter just the number)

Q1 (Spring):
Imagine you are a person that is 75kg, and takes up .09m³ of space. Now imagine you are on Jupiter where the surface gravity is 24.79 m/s².
What is your weight on Jupiter?
(note that some responses may be rounded)
75kg, 75N, 1860kg, 1860N, 830kg/m³

Q2 (F):
Imagine you are a person that is 75kg, and takes up .09m³ of space. Now imagine you are on Saturn where the surface gravity is 10.44m/s².
What is your mass on Saturn?
(Enter just the number)

Q2 (S):
Imagine you are a person that is 75kg, and takes up .09m³ of space. Now imagine you are on Saturn where the surface gravity is 10.44m/s².
What is your mass on Saturn?
(note that some responses may be rounded)
75kg, 75N, 780kg, 780N, 830kg/m³

Q3 (F):
Imagine you are a person that is 75kg, and takes up .09m³ of space. Now imagine you are on Neptune where the surface gravity is 11.15 m/s².
What is your density on Neptune?
(Enter just a number?)

Q3 (S):
Imagine you are a person that is 75kg, and takes up .09m³ of space. Now imagine you are on Neptune where the surface gravity is 11.15 m/s².
What is your density on Neptune?
(note that some responses may be rounded)
75kg, 75N, 830kg, 830N, 830kg/m³

Q4 (F):
Which do you expect to be the most dense?
Appendix A

The water in an Olympic sized swimming pool, the mercury in a mercury thermometer, all of Earth's atmosphere, they are all equally dense.

Q4 (S):
Which do you expect to be the most dense?
The water in an Olympic sized swimming pool, the mercury in a thermometer, Earth's atmosphere, they are all equally dense, there is no way to know.

Q5 (F):
Which do you expect to weigh the most?
The water in an Olympic sized swimming pool, the mercury in a mercury thermometer, all of Earth's atmosphere, they would all weigh the same.

Q5 (S):
Which do you expect to weigh the most?
The water in an Olympic sized swimming pool, the mercury in a thermometer, Earth's atmosphere, they are all equally dense, there is no way to know.
Standing Waves

*Please note the fall items were slightly out of order compared to how they are presented here

Q1 (F):
Which of these was closest to your fundamental frequency?
5Hz, 10Hz, 15Hz, 20Hz, 25Hz

Q1 (S):
Which of these was closest to the fundamental frequency of your standing wave?
7Hz, 14Hz, 21Hz, 28Hz, 35Hz

Q2 (F):
Particles of a material (sic) that move up and down perpendicular to the direction that the wave is moving are in what type of wave?
Torsional, longitudinal, mechanical, transverse, surface

Q2 (S):
Particles of a material that move up and down perpendicular to the direction that a wave is moving are in what type of wave?
Torsional, Longitudinal, Mechanical, Transverse, Surface

Q3 (F/S):
If the 6th harmonic of a particular standing wave is 78Hz, what is the 3rd harmonic?
234Hz, 26Hz, 75Hz, 39Hz, 13Hz

Q4 (F/S):
What is the amplitude of the above wave (image given)?
1, 2, 4, 8

Q5 (F/S):
The time required for one cycle of any repeating event is the:
amplitude., frequency., period., rotation. second.

Q6 (F/S):
If a wave travels 30 m in 1 s, making 60 vibrations per second, what are its frequency and speed, respectively?
30Hz and 60m/s, 30Hz and 30m/s, 60Hz and 30m/s, 60Hz and 15m/s, 15Hz and 30m/s

Q7 (F):
The mass per unit length (μ) of a uniform string changes depending on how much string there is.
True, False

Q7 (S):
The mass per unit length (linear density) of a uniform string changes depending on how much string there is.
True, False
Appendix A

Q8 (F):
Recall that $F = n^2LT$

If you increase $T$ by a factor of 4, and $L$ by a factor of 2, what will happen to $F$?
$F$ will increase by a factor of 4., $F$ will increase by a factor of 2., **$F$ will not increase or decrease.**, $F$ will decrease by a factor of 2., $F$ will decrease by a factor of 4.

Q8 (S):
The equation for the frequency of a standing wave is:
$F = n^2LT$

Where $F$ is the frequency, $n$ is the harmonic, $L$ is the length of the string, $T$ is the tension, and $\mu$ is the linear mass density.
If you increase $T$ by a factor of 4 and $L$ by a factor of 2, how does $F$ change?
$F$ will increase by a factor of 4., $F$ will increase by a factor of 2., **$F$ will not increase or decrease.**, $F$ will decrease by a factor of 2., $F$ will decrease by a factor of 4.
Simple Harmonic Motion

Q1 (F/S):
A simple harmonic oscillator oscillates with a frequency $f$ when its amplitude is $A$. What is the new frequency if the amplitude is tripled to $3A$?
$f/9$, $f/3$, $f$, $3f$, $9f$

Q2 (F/S): (This question was graded incorrectly in the fall, but was adjusted for analysis)
Simple pendulum A swings back and forth at twice the frequency of simple pendulum B. Which statement is correct?
**Pendulum B is $\frac{1}{4}$ as long as pendulum A.,** Pendulum A is twice as massive as pendulum B., Pendulum A is $\frac{1}{2}$ as long as pendulum B., Pendulum B is twice as massive as pendulum A., Pendulum A is $\frac{1}{4}$ as long as pendulum B.

Q3 (F/S):
What is the frequency when a weight on the end of a spring bobs up and down and completes one cycle every 2s?
$.5Hz$, $1Hz$, $2Hz$, $2.5Hz$, Depends on the mass

Q4 (F/S):
When the mass of a simple pendulum is quadrupled, how does the time $t$ required for one complete oscillation change?
Decreases to $\frac{1}{4} t$, Decreases to $\frac{3}{4} t$, Increases to $4t$, Decreases to $\frac{1}{2} t$, Stays the same

Q5 (F/S):
A floating leaf oscillates up and down two complete cycles each second as a water wave passes by. What is the wave's frequency?
$.5Hz$, $1Hz$, $2Hz$, $2.5Hz$, Depends on the mass

Q6 (F/S):
A simple pendulum consists of a mass $M$ attached to a weightless string of length $L$. Which statement about the frequency $f$ is accurate for this system when it is displaced a small amount from equilibrium?
The frequency is directly proportional to the period., The frequency is independent of the mass., The frequency is inversely proportional to the amplitude., The frequency is independent of the length., The frequency is dependent on the mass.

Q7 (F/S):
What is the period of an oscillator which oscillates at 100Hz?
$100s$, $50s$, $10s$, $.05s$, $.01s$
Generators

Questions were not presented in the same order in both quizzes.

Q1 (F/S):
Which of the following is the most similar to a gravitational field?
Electric charge, Electric current, Electric Field, Electromagnetic force, Electric dipole

Q2 (F/S):
If a loop of wire is rotating in a magnetic field, how must the loop be oriented in the field for there to be no current produced?
It must be at an angle of 45° with North., It must be perpendicular to the field., It must be at an angle of 45° to the South., It must be parallel to the field., There is not enough information to tell.

Q3 (F/S):
A wire is moved through a magnetic field and no current is produced. What can be concluded?
The wire is moving parallel to the field., The wire is moving perpendicular to the field., The gravitational field overpowers the magnetic field., There is no electric field outside the wire., The wire has no mass.

Q4 (F/S):
A loop of wire is rotated about its axis (which is perpendicular to a magnetic field). In one rotation, how many times does the induced current in the loop reverse directions?
It doesn’t reverse directions, Once, Twice, Four times, Ten times

Q5 (F/S):
Generators create energy.
Is the above statement true or false?
The above statement is true, because generators make energy., The above statement is true, because generators can make light and heat., The above statement is false, because generators create electricity., The above statement is false, because energy cannot be created., The above statement is false, because not all generators create energy.

Q6 (F/S):
Which of the following creates a current in a wire?
Rotating a wire outside a magnetic field., Rotating a magnet in a wire loop., Rotating a wire parallel to a magnetic field., Making a loop out of a wire., None of these.

Q7 (F/S):
When you hooked up the two hand generators to each other, what happens when you turn one of the handles through one full rotation?
The other handle makes less than one full rotation, The other handle makes one full rotation, The other handle makes more than one full rotation, The other handle does not rotate, None of these things happen

Q8 (F/S):
Where does the extra energy go if a generator is only 75% efficient?
The other 25% of the energy is destroyed, Nothing happens, the generator is just 75% efficient, The other 25% of the energy is lost to the surroundings, The other 25% of the energy doesn’t interact with the generator, None of these
Appendix A

Geometric Optics

Q1 (F/S):
As the angle of an incident (incoming) ray of light increases, the angle of the reflected ray: 
**Increases**, Decreases, Stays the same, Increases or decreases, Requires more information

Q2 (F/S):
If the index of refraction of diamond is 2.43, a given wavelength of light travels:
2.43 time faster in diamond than it does in air., 2.43 time faster in a vacuum than it does in a diamond., 2.43 time faster in a diamond than it does in a vacuum., 2.43 time faster in air than it does in diamond., 2.43 time faster in air than it does in vacuum.

Q3 (F/S):
Fill in the blank:
The angle of incidence is ____ the angle of refraction.
Equal to, greater than, less than; **greater than, less than, or equal to**; independent of

Q4 (F/S):
Let n1 be the index of refraction of the incident medium, and n2 be the index of refraction of the refracting medium. Which of the following must be true if the angle that the refracted ray makes with the boundary (not with the normal) is less than the angle that the incident ray makes with the boundary?
n1 < n2, n1> n2, n1 < 1, n2 < 1, n1 = n2

Q5 (F/S):
If a single lens forms a virtual image of an object, then the:
Image must be inverted, image must be larger, **lens could be either a diverging or converging lens**, lens must be a converging lens, lens must be a diverging lens

Q6 (F/S):
A lens forms a virtual image of an object. Which of the following must be true of the image?
It is larger than the object and upright., It is smaller than the object and inverted., It is the same size as the object and upright., It is inverted., **It is upright.**

Q7 (F/S):
Which statement about thin lenses is correct when considering only a single lens?
**A diverging lens always produces a virtual upright image.**, A diverging lens always produces a real upright image., A diverging lens always produces a virtual inverted image., A diverging lens always produces a real inverted image., A converging lens always produces a real inverted image.

Q8 (F/S):
The index of refraction is based on the ratio of the speed of light in:
Two different transparent materials, a solid to the speed of light in the transparent material, water to the speed of light in the transparent material, air to the speed of light in the transparent material., a vacuum to the speed of light in the transparent material.
Q9 (F/S): Except for air, the refractive index of all transparent materials is:
equal to 1, less than or equal to one, less than 1, \textbf{greater than 1}, greater than or equal to 1

Q10 (F/S): In the figure a ray in glass arrives at the glass-water interface at an angle of 48° with the normal. The refracted ray makes an angle of 68° with the normal. If another ray in the glass makes an angle of 29° with the normal, what is the angle of refraction in the water? (Use the index of refraction of water $n = 1.33$)

$29°, 37°, 31°, 41°, 46°$
Physical Optics

Question 3 occurred at the end of the assessment in spring.
Q1 (F/S):
Dark fringes are also called:
**Minima**, Maxima, Zero Points, Peaks, None of the above

Q2 (F/S):
In the double slit experiment, you measure the distance from the center of the pattern to the nth _____.
Dark spot, minimum, fringe, **bright spot**, none of the above are correct.

Q3 (F):
Which other lab this semester had a core concept that is also important for light diffraction?
Lab 1: Uniform Motion, Lab 3: Exponential Decay, Lab 5: Oscillations and Simple Harmonic Motion, **Lab 6: Sound**, Lab 7: Magnets

Q3 (S):
Which of the following physics phenomena is most important for diffraction?
Refraction, **Interference**, Reflection, Equilibrium, Conservation of Energy

Q4 (F/S):
What color laser did we use?
Purple, Blue, Green, Yellow, **Red**

Q5 (F/S): (Note that students could not go back and revisit previous questions)
If we used a blue laser instead of a red laser in the double slit experiment, how would the diffraction pattern change? Blue lasers have smaller wavelengths than red lasers.
**Nothing, only the color would change.**; The diffraction pattern would get wider.; The diffraction pattern would **get smaller.**; The diffraction pattern would look like the single slit experiment.; The experiment would not work with a blue laser.

Q6 (F/S):
What do you expect to have a wider diffraction pattern, a hair or a piece of string? Assume that the string is thicker than the hair.
The string, because it is made up of many pieces.; The hair, because it has an oval cross section.; The string, because it is larger.; **The hair, because it is smaller.**; They will be the same, because of Babinet’s Principle.

Q7 (F/S):
If we made the slit spacing in the single slit experiment smaller than what we did in lab, what would happen to the diffraction pattern?
**It would be wider.**, It would be smaller., It would change color., It depends on how small the spacing is., Nothing would happen.

Q8 (F/S):
What is the fundamental idea behind Babinet’s Principle?
Light is a particle and a wave. Light interferes with itself when it passes through 2 slits that are close to each other. Light diffracts the same way around a solid object and through a hole of the same size. When passing through a hole, there is more constructive interference than destructive interference, so a diffraction pattern occurs. The single slit and double slit experiment are the same experiment, you just add another slit.

Q9 (F):
Which of the following best explains why light diffracts?
Babinet’s principle, Wave particle duality, Refraction, Reflection, Interference

Q9 (S):
Which of the following was described after the discovery that light diffracts?
Babinet’s principle, Wave particle duality, Refraction, Reflection, Interference

Q10 (F/S)
Why do lasers diffract when going through a slit, but not when traveling through a doorway?
Doorways do not interact with lasers. Doorways are much more common so lasers have learned how to move through them. The frequency of the laser is too small for the doorway to cause diffraction. Laser light moves too fast for a doorway to cause diffraction. The size of the doorway is much larger than the wavelength of the laser.
Appendix A

Atomic Spectra

Q1 (F/S):
What happens to an atom when it absorbs energy?
**The atom re-emits the energy as light.**, The atom re-emits the energy as heat., The extra energy decreases the speed of the electrons in their orbitals., The atom stores the energy as kinetic energy., The atom re-emits the energy as an electron.

Q2 (F/S):
Which of the following types of electromagnetic radiation has the highest energy per photon?
1. Microwave  2. Infrared  3. Ultraviolet
1 only, 2 only, 3 only, 1&2 only, 1&3 only

Q3 (F/S):
Which of the following types of electromagnetic radiation has the highest energy per photon?
1. Microwave  2. Infrared  3. Ultraviolet
1 only, 2 only, 3 only, 1&2 only, 1&3 only

Q4 (F/S):
Electromagnetic waves consist of:
particles of heat energy., high-frequency gravitational waves., compressions and rarefactions of electromagnetic pulses., low-frequency gravitational waves., oscillating electric and magnetic fields.

Q5 (F/S):
Which form of electromagnetic radiation has photons with the lowest energy?
X-rays, Ultraviolet radiation, Radio Waves, Microwaves, Infrared radiation

Q6 (F/S):
Which of the statements about light is FALSE?
A packet of light energy is known as a photon., Color can be used to determine the approximate energy of visible light., Light travels through space at a speed of 3.0X10^8 m/s., Ultraviolet light cannot be seen with the unaided eye., **All statements are true.**

Q7 (F/S):
Which color of the visible spectrum has the shortest wavelength?
Violet, Green, Orange, Blue, Yellow

Q8 (F/S):
When neon light passes through a prism (like a diffraction grating). what is observed?
White light, Continuous spectrum, The same neon light, **Bright spots or lines**, Both A and D.

Q9 (F/S):
Light with the lowest frequency detected by your eyes is perceived as:
Violet, Green, Yellow, Orange, **Red**

Q10 (F/S):
Which of the following does NOT produce a continuous spectrum:
The sun, A fire, **Neon signs**, Incandescent lights, None of the above
Appendix A

Photoelectric Effect

Q1 (F/S):
Which photon has more energy, one with a long wavelength, or one with a high frequency?
Long wavelength, **High frequency**, They would have the same energy, Wavelength and frequency aren’t related to energy, Photons don’t have energy

Q2 (F/S):
Suppose light strikes a metal and a stopping voltage of 1.3V is observed. How would the stopping voltage change if the same light strikes a metal with a **smaller** work function?
The stopping voltage would stay the same, **The stopping voltage would increase**, The stopping voltage would decrease, It depends on the metal, There would be no stopping voltage

Q3 (F/S):
Which color of light has the most energy?
Red, Orange, Yellow, Blue, **Purple**

Q4 (F/S):
Blue light is shined on a variety of different metals, and there stopping voltages are given below, which metal has the largest work function?
Metal A: 2.3V, Metal B: 1.2V, Metal C: 5V, **Metal D: 0V**, Metal E: .5V

Q5 (F/S):
Which of the following is a good example of the photoelectric effect?
A solar sail, which is pushed through space by the light of a star., A light bulb, which turns electricity into light., The sun, which produces light with nuclear reactions., **A solar panel, which takes light and turns it into electricity.**, None of the above are good examples of the photoelectric effect.

Q6 (F/S):
Which color of light has the least energy?
**Red**, Orange, Yellow, Blue, Purple

Q7 (F/S):
The work function of a metal is a measure of:
How much work it takes to bend a piece of metal., How much you can bend a metal before it breaks., How hot you can heat a metal before it melts. How much light you have to shine on a metal before electrons fly off.. **How much energy you have to shine on a metal before electrons fly off.**

Q8 (F/S):
How does the intensity of the light shining on a metal influence the photoelectric effect?
A larger intensity increases the stopping voltage., A larger intensity decreases the stopping voltage., **A larger intensity increases the number of electrons removed.**, A larger intensity decreases the number of electrons removed., The intensity has no effect on the photoelectric effect.
Appendix A

Q9 (F/S):
How does the frequency of the light shining on a metal influence the photo electric effect?
**Higher frequency light increases the stopping voltage**, Higher frequency light decreases the stopping voltage, High frequency light increases the number of electrons removed, High frequency light decreases the number of electrons removed., The frequency has no effect on the photoelectric effect.

Q10 (F/S):
What color of light was not used in this week's experiment?
Red, **Orange**, Yellow, Green, Blue
Appendix A

Absolute Zero

Q1 (F/S):
Which of these is an absolute temperature scale?
Fahrenheit, Celsius, **Kelvin**, Rankine, None of the above

Q2 (F/S):
Which of the following relates to 0 Kelvin?
0K is the freezing point of water., **0K is when molecular motion stops.**, 0K is the temperature of the coldest place on Earth., 0K is the temperature of the coldest point in space., 0K is the freezing point of carbon dioxide.

Q3 (F/S):
Which scale has the larger temperature unit, Celsius or Kelvin?
Celsius has the larger temperature unit, Kelvin has the larger temperature unit, Niether Celsius or Kelvin has the larger temperature unit, They have the same temperature unit, It isn’t possible to relate the two

Q4 (F):
Which of the following had the lowest temperature?
Boiling water, Room temperature, Ice water, **Liquid nitrogen**, They all had the same temperature

Q4 (S):
Which of the following had the lowest temperature recorded in Thursday's experiment?
Boiling water, Room temperature, Ice water, **Liquid nitrogen**, They all had the same temperature

Q5 (F/S):
Which of the following is closest to room temperature?
0°C, 10°C, **20°C**, 30°C, 40°C

Q6 (F/S):
If you have an ideal gas with a fixed volume and a fixed number of moles, how will the pressure change if the temperature increases?
The pressure will increase., The pressure will decrease., The pressure will not change., It is not possible to tell without knowing the volume., It is not possible to tell without knowing the number of moles.

Q7 (F/S):
If you have an ideal gas with a fixed volume and a fixed number of moles, how will the temperature change if the pressure decreases?
The temperature will increase., **The temperature will decrease.**, The temperature will stay the same., It is not possible to tell without knowing the volume., It is not possible to tell without knowing the number of moles.

Q8 (F/S):
If the freezing point of water is 32°F, and the boiling point is 212°F, which temperature scale has a larger unit of temperature if the freezing point of water is 0°C and the boiling point is 100°C?
Appendix A

They have the same unit of temperature., Fahrenheit has the larger unit of temperature., **Celsius has the larger unit of temperature.**, You need more information to tell. You cannot compare these two scales because they measure different things.

Q9 (F/S):
(Answer, kPA in F, PSI in S; students were likely confused by inconsistencies between what the instructor said and what the lab manual had written, there was a last-minute equipment change.)

What pressure unit did you measure in for this experiment?
Torr, mmHg, PSI, atm, kPa

Q10 (F/S):
Which object has a higher temperature, one with a temperature of absolute zero on the Celsius scale, or one with a temperature of absolute zero on the Fahrenheit scale?
The object measured on the Fahrenheit scale. The object measured on the Celsius scale **The objects have the same temperature, but the scale readings are different.**, The objects have the same temperature, and the scale readings are the same., The objects don’t have the same temperature, but the scale readings are the same.
Appendix B

Appendix B: Demographic questions

25. What is your year in school?
   (a) First year of college
   (b) Second year of college
   (c) Third year of college
   (d) Fourth year of college
   (e) Other:

26. What is your gender identity?
   (a) Female
   (b) Male
   (c) Other
   (d) Prefer not to answer

27. Do you consider yourself an international student?
   (a) Yes
   (b) No
   (c) Prefer not to answer

28. What race or ethnicity do you consider yourself?
   a. White
   b. Hispanic or Latino
   c. Black or African American
   d. Asian
   e. Other

29. How many physics courses have you taken before this course (including high school)?
   (a) One
   (b) Two
   (c) Three
   (d) Four
   (e) More than four
30. How many laboratory courses (or courses with a lab component) have you taken before this course (including high school)?
(a) One
(b) Two
(c) Three
(d) Four
(e) More than four

31. I would like to study physics in order to satisfy
(a) Mainly my own interests and less what certain people expect of me.
(b) More my own interests than what certain people expect of me.
(c) My own interests and what certain people expect of me equally.
(d) More what certain people expect of me than my own interests.
(e) Mainly what certain people expect of me and less my own interests.
Appendix C

Appendix C: Comparison of pre-lab forms

Fall 2016 general pre-lab form

Phy103: Concepts in Physics Laboratory
Pre-Lab Questions
Lab #:

What is the stated goal of the experiment?

What raw data will you be collecting?

What will you calculate using the raw data?

Choose 3 major terms from the lab introduction and define them.

List one question that you have about the lab after reading the write-up. If you didn't find anything confusing, what do you think the hardest topic was for the most of your classmates? Why?
Phy103: Concepts in Physics Laboratory
Pre-lab Questions
Lab # 10 - Atomic Spectra

Name:
Lab Section:
Table:

= What is the goal of this experiment?

= What is an atomic spectrum and why is it called an optical fingerprint of a material?

= What is the name of the visible emission spectrum produced by Hydrogen?

= How do atoms produce light?

= What will we be using as a blackbody radiation source?

= List one question that you have about the lab after reading the write-up. If you didn't find anything confusing, what do you think the hardest topic was for the most of your classmates? Why?
Appendix D: Comparison of lab assignments

Fall 2016 Geometric Optics Lab

Lab 9: Geometric Optics

Group Members:

Section Letter: Table #:

Questions:

Ray Tracing:
1. Let the laser enter the flat side of your Plexiglas semi-circle. Rotate the laser and observe the behavior of the light as a function of incident angle. Record one example on a sheet of paper of input angle versus angle of propagation inside the Plexiglas (Don’t use 90° as the incident angle).

2. Calculate the index of refraction, n, of the Plexiglas using the measurements made above and Snell’s Law. Show your work below.

3. Calculate the angles of refraction that occur as a laser passes through a Plexiglas semi-circle, knowing that $n_1 = 1.00$, and $n_2 = 1.33$. On the back of the piece of paper from 1, draw what the path of the ray as it enters the semi-circle, goes through the semi-circle, and as exits the semi-circle. You chose the initial angle of incidence, but make sure it enters at some angle other than 90° and is off from the center.

Focal Lengths:
Place the large Plexiglas plano-convex lens (semicircle) on the paper. Trace its outline on the large paper found on your table. Shine the multi-laser source on the flat side of the lens. On the provided paper, trace the 5 beams as they enter the lens (be sure they are perpendicular to the surface), their paths through the lens and their paths outside the lens, until the rays intersect at the focal point. Label this point F on your sketch. For this experiment the focal length of the lens is the distance from the incident light (in this case the flat part) to the focal point F.
To determine the focal length of the plano-concave lens (the other one), you must extend the traced exit rays backwards through the position of the lens until they intersect and then measure the distance from the flat side of the lens to the intersection point for the focal length. This length is recorded as a negative distance.

4. What is that focal length $f$ of the plano-convex lens?

5. What is the focal length of the plano-concave lens?

6. What type of lens is your eye? Make sure to explain your reasoning in terms of what you learned in lab today. You can draw a picture on one of your sheets of paper if that will help you explain.

**Imaging:**
Take your two glass lenses and look at something small and nearby, maybe some words in a book or on the computer screen. Describe what you see below. How does this view change as you move the lens toward and away from your eye? Which type of lens causes the text to be magnified? Demagnified?

7. Describe what you see when you use the biconvex lens. Does the image ever flip? Does the image get bigger or smaller when you move the lens up and down? When is the image clear?

8. Describe what you see when you use the biconcave lens. Does the image ever flip? Does the image get bigger or smaller when you move the lens up and down? When is the image clear?

9. Use your lenses to form a clear image of the light in the ceiling on the table, which one lens works? What is the focal length of the lens (measure from the lens to the table)

10. Using the lens that made a clear image in 9, shine the parallel rays from the Multiple Laser Ray Source through the lens and see where they converge to a point on a piece of paper. Measure this focal length.
11. How do the values for focal length between 9 and 10 change? Which one would be more accurate, and why?

**Image projection:**
We want to determine the image distance and magnification of an image on a screen. To do this we will use the Gauss' lens formula and the setup from the lab introduction.

To get an image on the screen, move the lens back and forth until the images is clear and in focus on the screen.

12. Measure $d_o$ and $d_i$ when you get a clear image on the screen.

13. Using the measured image distance from question 12, and the focal length you chose in question 11, calculate $d_o$ using Gauss' lens formula, and compare the calculated value to the measured value from 12. Explain any differences.

14. You may have a magnified or a reduced image. Measure your image's magnification by measuring the length of the vertical arrow on the screen and dividing that length by the length of the vertical arrow on the mask. Record the length values and the value for magnification here.

15. Calculate the magnification of the image using your values for image distance and object distance using the equation for magnification from the introduction. Did each formula give the same value of $M$? What do you think the negative sign implies?

16. What should happen if you move the object a bit closer to or further from the lens? Try it and describe the result. Adjust the distance of the screen to keep a clear image.

17. How did $M$ and $d_o$ and $d_i$ change each time? Were the changes consistent with Gauss' lens formula and the formula for $M$?
Spring 2017 Geometric Optics Lab
(A lab was removed earlier in the semester, leading to different lab numbers.)

Lab 8: Geometric Optics

Group Members: 

Section Letter: 

Table #: 

>>>>>>SPECIAL SAFETY NOTE<<<<<<
DO NOT LOOK DIRECTLY INTO LASER

Questions:

Answer the following questions using the semi-circle Plexiglas lens and a laser pointer found near your table. You'll turn in the piece of paper with your tracings and calculations.

1. Let the laser enter the flat side of your Plexiglas semi-circle. Rotate the laser and observe the behavior of the light as a function of incident angle. Record one example on a sheet of paper of input angle versus angle of propagation inside the Plexiglas (Don't use 90° as the incident angle).

2. Calculate the index of refraction, n, of the Plexiglas using the measurements made above and Snell's Law. Show your work on your paper, and record your answer here.

3. Ask your TA for an incident angle, and using the index of refraction you calculated above find the refracted angle. On the back of your piece of paper show your work for finding the refracted angle.

Answer the following questions using the Plexiglas lenses and a laser array. Record your work on the large pieces of paper found at the front of the room. You'll turn in this piece of paper as well.

4. Shine the laser array at the flat side of the plano-convex lens, trace the incident and refracted rays, then determine the focal length of the lens. Record the value here. It might be helpful to trace the outline of the lens with the flat edge near the edge of the paper.
5. Shine the laser array at the flat side of the plano-*concave* lens, trace the incident and refracted rays, then determine the focal length of the lens. Record the value here. It might be helpful to trace the outline of the lens with the flat edge near the middle of the paper.

6. When you feel confident in your ability to work with Snell’s law and estimate the path a laser will take through a medium, let your TA know you’re ready for the quiz at the front of the room. Attach any extra work done on paper to the other 2 papers you’ll turn in today.

Answer the following questions using the glass lenses and the appropriate laser.

7. Describe what you see when you use the biconvex lens: how does the image change when you move the lens to and from the object? When is the image clear?

8. Describe what you see when you use the biconcave lens: how does the image change when you move the lens to and from the object? When is the image clear?

9. Using the lens which can form a clear image of the light in the ceiling on the table, determine the focal length (measure from the lens to the table) and record it here.

10. Using the lens that made a clear image in 9, shine the laser array through the lens and see where the multiple lasers converge to a point on a piece of paper. Measure this focal length and record it here.

11. How do the values for focal length between 9 and 10 change? Which one do you think is more accurate? Explain your reasoning.
Answer the following questions using the optical rail setup at your table.

12. Create a clear image on your screen and show your TA when you get a chance. How does the image on the screen compare to the object making the image?

13. Make a clear image that is larger than the object, then record the distance from the object to the lens and the distance from the image to the lens here.

14. Make a clear image that is smaller than the object, then record the distance from the object to the lens and the distance from the image to the lens here.

15. If you only know the distance from the object to the lens, and the distance from the image to the lens, can you determine if the image is bigger or smaller than the object? If so, how? If not, why not?
Appendix E: IRB Application

Human Subjects Research: Application for Level 1 Screening and Approval

**HS Research Protocol**: All research projects, involving human subjects, regardless of funding source must submit an application to the Research Ethics and Integrity Office (MREI) and obtain approval from the appropriate authority prior to the initiation of the research. Some research requires IRB review and approval (Level 2) and some research will be deemed to qualify for exemption from review (Level 1).

Completed applications should be emailed as attachments to humansubjects@miamioh.edu. Questions about methods or the status of an application may also be sent to this email address.

**A. Personnel Information** *(online help)*:

<table>
<thead>
<tr>
<th>PRIMARY CONTACT (Corresponding Researcher)</th>
<th>* Faculty Advisor (required for Student PC's):</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name: Jennifer Blue</td>
<td>Name:</td>
</tr>
<tr>
<td>Department: Physics</td>
<td>Department:</td>
</tr>
<tr>
<td>Phone: (513) - 529 - 1380</td>
<td>Phone:</td>
</tr>
<tr>
<td>Email address: <a href="mailto:bluejm@miamioh.edu">bluejm@miamioh.edu</a></td>
<td>Email address:</td>
</tr>
<tr>
<td>CITI Online Training (date): 9/2/2015</td>
<td>CITI Online Training (date): Date</td>
</tr>
<tr>
<td>MU Application training (date): 9/11/2002</td>
<td>MU Appl. training (date): Date</td>
</tr>
</tbody>
</table>

Primary Contact Status: □ Undergraduate Student □ Graduate Student ▼ Faculty/Staff

**Other Personnel**: persons who will interact with subjects or data that include explicit or implicit identifiers must be listed on a Personnel Form submitted with the application.

Note: All investigators and personnel that will be in contact with subjects or data that can be linked to subjects must complete initial online human subjects research training and then refresher training every three years. Please check to be sure all training is up-to-date before submitting your application: Training Completion Dates

**B. Project Information** *(online help)*:

<table>
<thead>
<tr>
<th>Application Title: Assessing the effect of inquiry labs on student content knowledge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Your Research Project Title: Assessing the effect of inquiry labs on student content knowledge</td>
</tr>
<tr>
<td>Human Subjects Research Dates: Anticipated projected beginning date: <strong>Sep 01 16</strong> (e.g. 1 Jan 17) (and after MREI approval)</td>
</tr>
<tr>
<td>Funding Sources if Federal, State, or otherwise external to MU:</td>
</tr>
</tbody>
</table>

71
C. RESEARCH OVERVIEW ([online help](#)):  

All subjects from whom data will be collected or consent obtained will be over 18 years of age: ☑ Yes ☐  

No The research activities of the subjects < 18 will occur in a normal classroom environment and involve normal classroom curriculum (i.e. research metrics include analysis of data that would be produced in the absence of research, e.g. test scores etc.)  

If “No” Above ☐ Yes (required to be normal curriculum)  

Note: If you cannot check “Yes” to either of the above descriptors, then you should be using the Level 2 application form.  

<table>
<thead>
<tr>
<th>Data Collection Methods - Research activities include:</th>
<th>☑ Survey Instruments (online or paper where the researcher does not have direct interactions with subjects)</th>
<th>☐ Interviews (in person, online, and or by phone)</th>
<th>☐ Observations of public behavior</th>
<th>☐ Data is publicly available (does not require a data use agreement with owner).</th>
<th>☐ Miami researchers have never or will never have access to data with subject identifying information (original data may have identifiers when collected outside of Miami but it must be provided to Miami researchers without identifiers).</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data is about:</td>
<td>☐ Personal facts about or experiences of the subject (beyond simple demographics).</td>
<td>☑ Effectiveness of a program or course.</td>
<td>☐ Factual information about a program, institution, or agency.</td>
<td>☐ Opinions about an institution (e.g. employer, school)</td>
<td>☐ Opinions about an issue (e.g. politics).</td>
</tr>
<tr>
<td>Confidentiality:</td>
<td>☑ Participation will include direct identifiers or identity could be readily deduced (if so, use of this form requires that the information collected pose a minimal risk to subjects).</td>
<td>☐ Participation will be anonymous without potential identification (e.g. an online survey where the IP address for a computer/lab could not easily be linked to subject identity and insufficient demographic information that could be linked to identity)</td>
<td>☐ Information will be recorded confidentially in such a manner that subjects could not be identified (directly or through information that could identify the subjects), e.g. an interview will be conducted but notes will not include identity and the interview will not be audio/video recorded.</td>
<td>☐ Data may reasonably include personal or opinion information that if accessed by anyone outside the research team could be a source of harm or embarrassment to the subject (whether it is meant to be anonymous or confidential).</td>
<td></td>
</tr>
<tr>
<td>Risk:</td>
<td>☐ Disclosure of the subjects responses outside the research team might place the subjects at risk of criminal or civil liability or be damaging to the subjects’ social or financial standing, employability, or reputation. This includes disclosure to peers and</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>family. If checked, data security should be addressed in section D(i) below.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>✗ Disclosure of the subjects responses outside the research team reasonably <strong>would</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>not place the subjects at risk of criminal or civil liability or be damaging to the subjects’</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>social or financial standing, employability, or reputation. This includes disclosure to</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>peers and family.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
D. Research Description (Complete all sections. If you believe a section does not apply to your research, you must explain and justify this or reference an attachment; Refer to the guidance document for help). (online help)

i) Purpose, Population and Methods: Use the space below to briefly describe the purpose, nature of the research, e methods to be used, and describe the subject population (demographics and approximate number of subjects). If necessary for more complex projects, descriptions can be expanded upon in appendices.

General Research question/objectives (online help):

The object of this research is to look at the effects of direct teaching of scientific reasoning on student content knowledge in a lab setting. There is little available research on the relationship between reasoning and content knowledge, and data on this topic is especially rare concerning lab classes. The experiment is occurring while the department is improving the quality of their non-majors lab class, so this experiment will be studying the changes made. Ideally, this study would show that there is a link between student’s ability to reason and their understanding of lab material.

Research Subject Population (online help):

The students enrolled in PHY103 are the target subject population for this study. In previous classes the population is mostly international students and largely male. There seems to be a mix of underclassmen and upperclassmen in the course. To ensure that only eligible individuals are studied, there will be a consent form given out at the beginning of the course outlining what data will be used for research and giving the students a chance to opt out if they are underage or do not wish to have their scores used in the research.

Estimated/Target number of subjects 200

Research Methods (online help):

To attempt to measure the effect of planned laboratory changes, a few additional surveys will be given to the students in addition to their regular coursework. At or near the beginning of the course a consent form explaining the research will be given along with a demographic survey and the Lawson Classroom Test of Scientific Reasoning; the Lawson test will be given again at the end of the semester. The consent form and the demographic survey will be given on paper, and the Lawson test will be available on paper or online. These 3 surveys, the demographic survey and the pre-post Lawson test, will be the only instruments used that are outside the purview of what would usually take place in the course. Quiz scores and final grades will also be used to assess student content knowledge.

Confidentiality/security methods:
Once students have filled out consent forms, students who do not wish to have their data used will have their data removed from the study. At the end of each semester, data from quizzes and surveys will be matched if possible, and then all identifying information will be removed. Results from the data will be presented in aggregate form, and they will be published in a thesis and possibly a journal.

ii) **Consent Process:** Use the space below to briefly describe the consent process that will be implemented. It is most expedient to compose a draft of your consent form and attach it to your application submission message or copy-paste it here.

**Invitation content and method (s):** e.g. social media posting text, in-person script, etc. ([online help](#))

Students will be given a consent form on paper at the beginning of the class, there will also be information on a copy of the form they take with them and on the Canvas site which informs them that they are able to discontinue participation at any point. Students will be informed that there are no repercussions for opting out of the research, and they can still complete the research surveys for extra credit if they wish to.

**Consent Presentation:** i.e. how consent information will be presented to the potential subject/parents. ([online help](#))

There will be a research consent form at the beginning of the course allowing researchers to study the average grades and individual assignments in the course. In addition to the regular course content to PHY103, there will be a few opportunities for extra credit to help with a research project. At the beginning and end of the course there will be a survey of scientific reasoning given which can be completed for extra credit. This survey is completely optional. There will be a demographic survey given alongside the consent form which is also completely optional. At the end of the semester all identifying information will be removed from the collected data, so your scores and survey results will remain confidential. If at any point during the semester you wish to discontinue participation in the research project, please contact your TA or the professor assigned to the class, and they will inform the research team.

**Consent Information** ([online help](#) and sample templates):

The consent form is attached within the appendicies, and is labeled bluejm_02_consent form

iii) **Appendices:** Use the space below to indicate the materials/documents that are included in the application. Please name the files in the order they are referenced in the description above or combine them into a single document (e.g. pdf, Word, or RTF) to submit as an attachment to the application message. ([online help](#))

- [x] Additional Personnel Form
- [ ] Invitations
Appendix E

- Consent Form(s) or Information Sheets – subject or parental consent
- Assent (Forms) or Information Sheets – for subjects < 18 years of age
- Survey Questions (for online surveys, we prefer to have questions, not screen shots of webpages)
- Baseline Interview questions to include demographics recorded (whether in person, online, or by phone)
- Other (list)

  Example quiz questions
  Lawson Classroom Test of Scientific Reasoning

E. INVESTIGATOR’S ASSURANCE STATEMENT

1. I agree to accept responsibility for the ethical conduct of research conducted in this project;
2. I agree to obtain approval from the Miami Research Ethics and Integrity Office prior to significantly modifying any of the procedures;
3. I attest that the information submitted in this application is true to the best of my knowledge.

Students listed as the PC must send their applications and appendices to their faculty advisor for the project using their institutional email account. The faculty advisor will endorse and send the application to humansubjects@miamioh.edu using your institutional email account. Use of your institutional email account serves as your signature and pledge to abide by the conditions stated above. This cover page document should be sent in its current format (Word Doc), no original signatures are required.

Primary Investigator Signature  _____ (Submitting by MU email serves as your signature) _____ Date: mmm dd yy

Staff Advisor Signature  _____ (Submitting by MU email serves as your signature) _____ Date:
Appendix E

Consent Form

Research Consent Form

Dear Student:

Jennifer Blue, a professor in the Department of Physics, and the Physics Education Research Group in the department are conducting a research project, and are asking for your help. We are interested in studying the department’s improvements in the PHY103 laboratory, and you have a unique perspective.

We invite you to participate in our research study. Involvement will include:
- providing some simple demographic information (this will be the only activity that is not a normal class assignment).
- researchers linking the demographic information to your grades (assignments and final grades). No person other than your Professor or lab teaching assistants will have access to your grades when identifiers are included. Data as with all course data will be treated as confidential.
- You will have the option to withdraw from the research at the end of the semester after final grades have been assigned.
- Identifiers will be removed from the research data 30 days after the end of the semester. The de-identified data will be used for analysis and secure long-term storage.

The generalized results may be presented at professional conferences or published in articles describing the results of the research. Data will not be presented in any way that individuals can be identified.

If you have further questions about the study, please contact Jennifer Blue at 513-529-1380 or bluejm@miamioh.edu. If you have questions about your rights as a research participant or the voluntariness of this consent process, please call the Office of Advancement of Research and Scholarship at 529-3600 or email: humansubjects@miamioh.edu. You may also register your concerns with the Physics Department Chair: Dr. Herbert Jaeger.

Thank you for your participation. We are very grateful for your help. Your signature below indicates you are over 18 years of age and are volunteering to participate in this research.

Again, thank you for your help.

Yours truly,

Jennifer Blue and the Physics Education Research Group
Department of Physics
Miami University

________________________________________________________
Name (printed)

_______________________________________________Date ___________
Signature

The Family Education Rights and Privacy Act (FERPA) is a Federal law that protects the privacy of student education records, both financial and academic. For the student’s protection, FERPA limits release of student record information without the student’s explicit written consent. By signing this form you are consenting to the release of information as described in the research description above. Note: once identifiers have been removed from the data and the data is added to the data set to be analyzed, it will be impossible to extract your individual data if you withdraw from the research project.