Equity and Introductory College Physics Labs

by

Danny Doucette

B.Sc., Saint Mary's University, 2006

M.Ed., Montana State University, 2011

M.A., University of Latvia, 2017

M.Sc., University of Pittsburgh, 2019

Submitted to the Graduate Faculty of

the Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2021

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Danny Doucette

It was defended on

May 27, 2021

and approved by

Dr. Russell Clark, Senior Lecturer, Department of Physics and Astronomy

Dr. Adam Leibovich, Professor, Department of Physics and Astronomy

Dr. Robert Devaty, Associate Professor, Department of Physics and Astronomy

Dr. Kathy Koenig, Professor, Department of Physics, University of Cincinnati

Thesis Advisor/Dissertation Director: Dr. Chandralekha Singh, Professor, Department of Physics and Astronomy

Equity and Introductory College Physics Labs

Danny Doucette, PhD

University of Pittsburgh, 2021

Labs have long been considered an important component of physics education. However, while there has been progress in transforming physics education in physics lecture classes to improve student learning, less progress has been made in introductory physics labs. At the same time, the physics community has come to recognize that there is a need to transform instruction in order to improve equity and inclusion in the discipline. Inequities may be manifested in the curriculum, the instruction or, especially, interactions between students. In physics labs, where students typically work collaboratively in groups of 2-4 students, some of these inequities may be especially salient.

This dissertation addresses the question of equity in introductory physics labs by considering student interactions, curriculum, and professional development for lab instructors. I report on qualitative studies that sought to outline mechanisms by which inequities are perpetuated and exacerbated in introductory labs, such as gendered task division, isolation, and stereotype threat. I conducted quantitative analyses to assess the impact of task division and diversity on student performance. I performed case studies of two lab transformations, shedding light on what works, and what slips through the cracks, when lab instruction is transformed. I analyzed the impact of transformed labs on student attitudes toward experimental physics, and I advance a framework for designing lab curricula and lessons to improve student learning. I developed and analyzed an approach to providing professional development for graduate student teaching assistants (TAs) who served as lab instructors. Finally, I analyzed student reflections about the impact of the transition to online learning.

Throughout this dissertation, I have focused on uncovering pedagogical advice for lab instructors. Some take-away messages include the importance of designing lab-work that requires students to share, not split, collaborative work; the need to avoid forming groups with isolated women students; the value of designing learning experiences with an expansive framing; and the effectiveness of role-playing in training TAs. It is my hope that this body

of work may serve to help instructors transform physics lab learning to improve equity and inclusion in physics.

Table of Contents

List of Tables

List of Figures

1.0 Introduction

Laboratory work has long held a cherished place in physics education in the U.S.A. [64]. Yet, while the goals and pedagogy associated with introductory college physics labs have oscillated and evolved with time [31], systemic scholarship of student learning and experiences in labs is relatively new [136]. This dissertation seeks to contribute to scholarship on student learning in introductory college physics labs by addressing three core aspects: student interactions, curriculum, and instruction.

1.1 Student Interactions

In introductory physics lab courses, students typically work in small groups of 2-4 to cooperatively complete a set of investigations. The requirement that students collaborate while engaged in learning activities within the culturally-rich domain of physics means that interactions between students who are working together can have a significant impact on student experiences and learning. These interactions can be especially important when groups are composed of students with diverse identities. Past research has documented significant challenges when students collaborate in physics with peers with different identities [24, 151, 246, 260]. In this dissertation, I focus on student interactions within the physics lab using both qualitative and quantitative analyses to extend this research.

Although I use a variety of methodological and interpretive frameworks in studying student interactions, two aspects of the framework are especially handy. First, I frequently rely on a communities of practice framework when analyzing interactions of students who are working in small groups. In a community of practice, participants form a small community that engages together in a set of common practices in order to complete a larger mission within the domain of their work [300, 301]. Specifically, I study how students act as a small community when they work together to complete their lab-work, and seek to understand how the practices they employ while working together might advance their work and impact their peers.

Second, since much of the analysis deals with students' identities, it is important to be clear about how I view students' identities impacting their experiences in the physics lab. I adopt the view that gender, race/ethnicity, and other external aspects of identity manifest in interpersonal interactions primarily through performance. In the ways that matter, gender and race are things that we do [115, 217, 302], not biological essentials. Thus, for example, when I view students with different genders having different experiences in the lab, I view that difference as arising from students navigating how they 'do gender' and 'do physics' differently.

1.2 Curriculum

Much scholarship in the past decades has focused on the relative benefits of evidence-based active engagement, a pedagogical approach that is believed to improve average student learning [99]. The context of the work in this dissertation is a set of introductory college physics labs that transitioned from a traditional, highly structured format to an evidence-based active learning approach. During this transition, the RealTime Physics curriculum was used. In RealTime Physics, students make predictions, conduct simple experiments with accessible apparatus to check their predictions, and engage in practice with concepts using multiple representations [278]. Additional changes to the curriculum included adding 'checkpoints' for students to check with their instructor while doing the lab, creating digital lab reports that streamlined the RealTime Physics assignments while also adding reflection questions about the nature of science, and introducing collaborative activities for remote group-work via Zoom.

One approach that I used in developing curricular materials for the introductory college physics lab was expansive framing. Expansive framing calls for instruction to make explicit connections between different contexts, and to both scaffold and position student work in a broader context, so that students are more easily able to transfer their learning to new contexts [93].

In order to quantitatively assess the impact of curricula on student learning, I used two types of assessments. First, the E-CLASS 'expectations and epistemology' [308] survey provides a measure of students' attitudes toward different aspects of experimental physics. Second, concept inventories like the FCI [130] gauged student understanding of physics concepts.

1.3 Instruction

Instruction is also an important factor that impacts student learning and experiences in educational settings. Many of the introductory college physics labs that served as the site for much of this research were taught by graduate student teaching assistants (TAs). The honors physics lab was taught by two faculty members, and the introductory physics courses that are required as pre-requisites for the lab course are taught by a rotating pool of faculty.

In this dissertation, I share two efforts to improve the effectiveness of instruction. I developed professional development for TAs that includes role-play, discussion, and reflection during weekly lab TA meetings. With the switch to online learning in 2020 because of the pandemic, I surveyed students and provided a summary of the affordances and constraints they perceived in their online physics classes.

Note that, while this dissertation was written by a single person, the research reported below was conducted by teams. Therefore, the pronoun 'we' will be used henceforward.

1.4 Overview

In Chapter 2, we used interviews and reflexive ethnographic observations to identify and analyze two common modes of work that may disadvantage female students in introductory physics labs. Students who adopt the Secretary archetype are relegated to recording and analyzing data, and thus may miss out on much of the opportunity to grow their physics and science identities by engaging fully in the experimental work. Meanwhile, students in the Hermione archetype shoulder a disproportionate amount of managerial work, and also may not get an adequate opportunity to engage with different aspects of the experimental work that is essential for helping them develop their physics and science identities. We use a physics identity framework to investigate how students under these modes of work may experience stunted growth in their physics and science identity trajectories in their physics lab course. This stunted growth can then perpetuate and reinforce societal stereotypes and biases about who does physics. Our categorization not only gives a vocabulary to discussions about equity in the physics lab, but may also serve as a useful touchstone for those who seek to center equity in efforts to transform physics instruction. This work previously appeared in [74].

In Chapter 3, we investigate the underrepresentation of women in physics by reflecting on interviews with two undergraduate women. Leia is a chemistry major who loves college-level physical chemistry and quantum mechanics but does not identify with the discipline of physics, partly because she has a low level of self-efficacy as a physicist and has received very little recognition for her work and learning in physics. Paulette is a physics major who loves physics but feels isolated by the current physics learning environment. She reluctantly dropped honors introductory lab after being snubbed by her male classmates who partnered with one another, leaving her to work alone. Paulette's experiences with condescending male professors activated a stereotype threat about who can succeed in physics that caused her to disengage in class. We also discuss what these women felt has helped them so far and explore their suggestions for what would help women in physics courses as they pursue their quest for a physical science degree. This work previously appeared in [79].

In Chapter 4, we report on an ethnographic study in which two researchers observed introductory physics labs. We found that many women in mixed gender groups adopted the role of group leader or project manager and ensured that the group stayed on task and completed the lab work as expected throughout the semester. Here we report on an investigation of the views about the physics lab of four such female pre-medical students with high agency who came across as group leaders in a traditionally-taught introductory physics lab course for bio-science majors, and who strived to ensure that their group did well in the lab. Our findings are based on semi-structured interviews with these students.

The interviews focused on diverse issues including the roles of their male lab partners, other peers and the teaching assistant in their learning in the physics lab, their views about learning physics in lecture and lab courses, the role of physics labs in promoting conceptual understanding, learning in physics lab compared with other science labs, and the role of bio-inspired labs in their learning. We find that these female student group leaders had surprisingly similar views about these issues pertaining to the physics lab. We recommend that departments trying to revamp their physics lab courses reflect upon these findings in order to make the labs more effective. This work previously appeared in [81].

In Chapter 5, we describe different types of collaboration in terms of the mutuality of engagement they represent, and investigate how these different types of collaboration impacted student interest and self-efficacy in introductory physics labs. We surveyed college physics students about their beliefs and experiences regarding working with a lab partner. We find that when asked explicitly about what they valued in a lab partner, a majority of students noted that they wanted a “fair split” of the work. However, we find that students experienced improved physics interest and self-efficacy when they participated equally in all aspects of the lab, such as operating the apparatus and recording the data, which is a different form of mutuality of engagement. This form of participation disproportionately benefited women, a traditionally disadvantaged group in physics labs. Our findings suggest that students’ physics interest and self-efficacy might be positively impacted in lab courses that are designed to ensure that students participate equally in all aspects of the lab work, as opposed to lab courses in which students split up their work inequitably, as they might prefer to do. This work previously appeared in [76].

In Chapter 6, we analyze collaborative group work, an important aspect of many introductory physics labs. This research focuses on how the gender diversity of a group of students impacts their engagement in the collaborative portion of an online introductory physics lab. We conducted a mixed methods study to determine the relationship between the gender composition of groups and the degree of engagement those groups demonstrated in weekly collaborative activity reflections, using randomly-assigned groups of four students. We also investigated the impact of giving students explicit grading rubrics. We conducted a series of semi-structured interviews with students in order to understand their perspective and experiences collaborating with others. When students were not provided with grading rubrics, we found that groups of all men had highest levels of engagement, and groups with only one woman had lowest levels of engagement as evidenced by the length of written reflections. On the other hand, when students were provided with explicit grading rubrics, groups with one or more women had higher, comparable, levels of engagement. In interviews, women described interactions in majority-male groups that were detrimental to effective collaboration, such as being ignored or talked over. These results suggest that instructors could improve equity in online lab courses (and perhaps other classes as well) by assembling student groups that do not have a single, isolated, woman group member or providing clear grading expectations.

In Chapter 7, we present a case study on the transformation of a physics lab. New technology like the Arduino microcontroller platform presents an opportunity to transform Beyond the First Year (BFY) physics labs to better prepare physics students for work in research labs and beyond. The flexibility, low cost, and power of these devices provides an attractive way for students to learn to use and master research-grade instrumentation. Therefore, we introduced new technology, including Arduino Due microcontroller boards, to a second-year honors physics lab in order to provide improved learning experiences for students. This transformation was implemented in three lab modules and focused on diminishing the black box nature of the traditional labs while encouraging students to engage in troubleshooting. The importance of troubleshooting was made evident to students by the instructor emphasizing it as an inevitable and central part of experimentation. This lab transformation also required that students perform work that was 'above and beyond' the scope of the assigned experimental work for part of the course credit. While the technological aspects of the transformation were received well by a majority of students, our observations during the initial implementation suggested a need for some modifications to instructional practices in order to improve the learning and experiences for all students. In particular, we find that many students can benefit from additional scaffolding in order to complete 'above and beyond' work. Similarly, we find that students in general, and underrepresented students such as women in particular, may need thoughtful intervention from the instructor, e.g., in order to avoid becoming isolated when the lab work is designed for pair work. Otherwise, some students may be left to work alone with a disproportionate work-load if students choose their own partners. With these lessons taken into account, recent student experiences in the transformed lab were notably improved. This work previously appeared in [78].

In Chapter 8, we begin to analyze the impact of the adoption of inquiry-based lab pedagogy on student attitudes toward experimental physics. There is a growing recognition of the need to replace “cookbook”-style introductory labs with more-meaningful learning experiences. To identify the strengths and weaknesses of a mix of cookbook-style and inquiry-based labs, an introductory lab course currently being reformed was observed following a reflexive ethnographic protocol and pre and post E-CLASS surveys were administered. We analyzed data to identify shortcomings of the current labs and to determine areas for improvement. This work previously appeared in [72]

In Chapter 9, we continue our analysis of the relationship between lab curricula and students attitudinal and conceptual understanding. Conceptual inquiry-based introductory physics lab curricula, such as RealTime Physics, have been shown to improve students' understanding of physics concepts, and therefore may be attractive for instructors who seek to transform their physics labs. However, the impact of conceptual inquiry-based lab instruction on students' attitudes and beliefs about experimental physics, as measured by the E-CLASS survey, is not yet fully understood. We present data from three curricular approaches over four semesters ($n = 701$). While we did not see improvement in E-CLASS scores in the first implementation of a conceptual inquiry-based introductory physics lab, the addition of questions that asked students to reflect on issues relating to experimental physics was associated with E-CLASS outcomes that are comparable to other effective approaches to lab instruction. These results suggest that conceptual inquiry-based lab curricula may be suitable for instructors seeking to transform introductory physics lab courses if student attitudes toward experimental physics are an important aspect of the transformation, provided that the curricula are supplemented with reflection questions.

In Chapter 10, we address the issue of curriculum and lesson design in introductory physics labs using a framework called expansive framing. Expansive framing, which positions students as participants in larger conversations that span time, place, people and disciplines, can be a valuable approach for designing curricula and learning experiences to help students learn physics through an interdisciplinary approach. This chapter reports on efforts to use expansive framing as a guiding principle while transforming and revitalizing an introductory physics laboratory class. In this chapter, we describe student experiences with two central elements of the lab course that were strongly influenced by the concept of expansive framing and related to interdisciplinary learning. First, we sought to incorporate and emphasize experiences related to the real-world and professional experiences of students, such as connections between biology and physics, that will be interesting for the bio-science majors interested in health-related professions who take the lab. Second, we sought to promote discussions between students and their graduate student instructors about the epistemology of experimental physics, which we refer to as the nature of science, which is an important interdisciplinary goal for the lab class. We explore the need, design, and implementation of these two elements of the lab course through analysis of student interviews and coursework. Consequently, we propose that using expansive framing for the design of student learning should be considered a best practice for the implementation of introductory college physics laboratory courses when seeking to adopt an interdisciplinary approach to student learning. This work will appear in [80].

In Chapter 11, we present and begin to evaluate a model providing professional development for graduate student teaching assistants (TAs). We developed and implemented a research-based professional development program that focuses on preparing TAs to effectively support inquiry-based learning in the lab. We identify positive effects by examining three possible ways in which the professional development might have impacted TAs and their work. First, we examine lab TAs' written reflections to understand the effect of the program on TAs' ways of thinking about student learning. Second, we observe and categorize TA-student interactions in the lab in order to investigate whether TA behaviors are changing after the professional development. Third, we examine students' attitudes toward experimental science and present one example case in which students' attitudes improve for those TAs who 'buy in' to the professional development. Our results suggest lab TA professional development may have a tangible positive impact on TA performance and student learning. This work previously appeared in [73].

In Chapter 12, we delve deeper into the model for providing professional development for TAs by adopting a framework of expectancy-value theory and cognitive apprenticeship. This chapter presents a specific instantiation of a model for lab TA professional development that uses a combination of cognitive apprenticeship and expectancy-value theories as its framework. We describe how our model was implemented in the lab TA professional development program, which included reflections, role-playing and other pedagogical activities offered through weekly meetings. Our evaluation included an analysis of TA writing and interactions with students alongside informal observations and interviews. We discuss the importance of accounting for TAs' interest and self-efficacy development in teaching the labs, as well as the challenge of motivating TAs who have very low initial levels of interest in supporting student learning. We find that many TAs in our lab TA professional development program demonstrated an improvement in TA performance in supporting student learning. Given that the professional development activities require only a modest investment of time, these positive results suggest that the model of lab TA professional development may be usefully adopted and adapted at other institutions where introductory labs are led by graduate student TAs. This work previously appeared in [75].

In Chapter 13, we turn to the switch to online instruction that happened because of the COVID-19 pandemic. In 2020, many instructors and students at colleges and universities were thrust into an unprecedented situation as a result of the COVID-19 pandemic disruptions: even though they typically engage in in-person teaching and learning in brick and mortar classrooms, remote instruction was the only possibility. Many instructors at our institution who had to switch from in-person to remote instruction without any notice earlier in the year worked extremely hard to design and teach online courses to support their students during the second half of 2020. Since different instructors chose different pedagogical approaches for remote instruction, students taking multiple remote classes simultaneously experienced a variety of instructional strategies. We present an analysis of students' perceptions of remote learning in their physics and other classes at a large research university in the USA, focusing on lecture-based, active learning, and lab classes; collaboration and communication; and assessment. Student reflections emphasized the importance of community and opportunities to study with peers; grade incentives; and frequent, low-stakes assessments. Reflecting on the challenges and successes of different types of remote instructional approaches from students' perspective could provide useful insight to guide the design of future online courses.

Finally, in Chapter 14, I discuss some potential future directions for this work.

2.0 Hermione and the Secretary

2.1 Introduction

> In lab, I'm usually in charge of writing down the data that we collect, and [my partner] is usually the one doing the physical part.

> I think my partners weren't always prepared for the labs, so it fell on me to understand and get the group to finish the lab... I need to be prepared to know what's going on, because they won't.

Consider the above quotes from students describing their experiences in introductory physics labs. Who do you imagine these students to be? How might students' genders affect the way they experience the traditional introductory physics lab?

The introductory physics lab presents a unique and powerful opportunity for students to grow their physics and science identities. Identity in this sense is the 'kind of person' [106] students consider themselves – with respect to physics, or with respect to science generally – and we may understand the lab as contributing to their larger physics or science identity trajectory. Well-designed labs can be particularly effective for identity growth because of their low-stakes nature, which allows students to 'tinker' with the apparatus and develop a meaningful and relevant understanding of physics as an experimental science, and because lab-work can be collaborative and engaging for students.

However, as physics lab instruction increasingly adopts pedagogical approaches that include evidence-based active engagement strategies [1, 28, 88, 101, 165, 169, 223, 275, 278] and collaborative learning [28, 52, 88, 125, 152, 173, 229, 224, 262, 267], concerns have emerged that these types of learning environments might actually increase the 'gender gap' even as all students are learning more than they would in traditionally-taught courses [161]. In particular, if physics lab environments are not equitable and inclusive, social interactions around physics may allow for activation of stereotype threats [197] and the perpetuation or verbalization of stereotypes about who belongs in physics and who is capable of succeeding in physics [124]. Additionally, in such an environment, micro-aggressions, discrimination, and harassment [13] have the potential to stunt the physics and science identity development of students from traditionally-disadvantaged groups if equity is not placed at the center of the learning process in designing the learning environment.

Likewise, research shows that due to lack of role models and societal stereotypes associated with physics, women in college physics classes report lower levels of self-efficacy [205], are more susceptible to stereotype threats [197], are more often subject to stereotypes related to their competence, and enroll as physics majors and graduate students at markedly lower rates [234] compared with their male peers. In introductory labs, women average less expert-like responses on E-CLASS [314], an assessment of student attitudes toward experimental physics, and may perform different roles when engaging in lab-work with male peers [141, 237]. Research also shows that as they progress in their careers, female graduate students and scientists continue to experience inequities in research labs [61, 115]. Thus, promoting positive physics and science identity development [4] by creating an equitable lab learning environment is especially important for students from traditionally-disadvantaged groups, including women, as we seek to rectify longstanding inequities in physics.

This research is concerned with how gender is expressed in an introductory physics lab if no explicit effort is made to create equitable and inclusive learning environments where all students thrive and how it may disadvantage some students. In particular, we analyze our observational and interview data from the lens of students in the introductory physics lab 'doing gender' [115, 302] while 'doing physics'. Thus, if physics is framed, presented, and conducted in ways and in learning environments that are more aligned with traditional conceptions of masculinity and femininity, students are likely to position themselves and perform in response to these conceptions and reconcile 'doing physics' with 'doing gender' [37, 115, 256]. In this research, it is in students' navigation of aspects of their gender while doing physics that we seek clues about how we can improve instructional practices, learning environments, and lab cultures to positively impact students' physics and science identity trajectories [61]. We note that we recognize that gender is not a binary construct, however, all students in this investigation voluntarily self-reported identifying as male or female.

With this lens in mind, the introductory physics lab is at a crossfire: required for a large portion of the student body in science and engineering, fundamentally collaborative, increasingly adopting active learning approaches, and largely unattuned to the impact it is having on the physics and science identities of traditionally-disadvantaged students such as women and racial and ethnic minority students. Unlike traditional lecture-style courses, both labs and reformed courses that use collaborative evidence-based active-learning approaches (such as flipped classes) may allow gender stereotypes about physics to become especially salient and relevant. However, physics learning environments should not be allowed to perpetuate negative stereotypes about who can do physics, and about who can develop a strong identity as a physics or science person. Instead, physics learning environments should help all students develop physics and science identities. To that end, this research may provide useful insight for both labs and courses that employ collaborative learning.

The goal of this research was to use reflexive ethnographic observations in introductory labs and individual interviews with students in those labs in order to identify and map out how traditional lab instruction may impact students who work in mixed-gender groups of two or three students in our traditional introductory labs. In these labs, run by graduate student TAs, there is typically no explicit effort made to make the learning environment equitable. We identify two modes of work in which women may be disadvantaged. In the Secretary-Tinkerer mode, men tend to monopolize tinkering with the apparatus while women tend to be found in a note-taking or supportive role. In the Hermione-Slacker mode (Hermione is named for the clever, devoted, hard-working character from the *Harry Potter* [100] series who exemplifies the role in contemporary media), women tend to be thrust into the role of managing the experimental work, communicating with peers and the instructor, preparing for the lab, and doing most of the work in each lab session while their partners make minimal contributions. Finally, we discuss some research-based approaches that may help to reduce the prevalence of these modes of work.

2.2 Framework

We employ an identity framework to analyze how introductory physics lab learning environments affect the development of physics and science identities for female and male students [44]. In this framework, physics identity pertains to whether students see themselves as a physics "kind of person" [106, 286]. We also acknowledge that a student's identity "is not predetermined and fixed" [44] and that one's identity is dynamic and "always being shaped and impacted by one's environment" [148]. An identity framework is ideally-suited to the analysis of students' experiences in culturally-rich settings [291, 120] such as the introductory lab because identity framing focuses on and values the experiences of individual students, while avoiding the trap of deficit models that may be interpreted as inadequacies from differences between students. In our case, we seek to understand whether the way that physics lab learning environments are designed ensure that all students develop a stronger identity as a physics or science person.

Three constructs are often discussed in connection with physics identity. Perceived recognition is the degree to which students feel recognized or valued by peers, TAs, instructors, and family as a physics person or a person who is good at physics. Research suggests both that recognition is the strongest influence on the development of physics identity, and that the average perceived recognition by the instructor/teaching assistant in physics courses is larger for men than for women [155]. Interest is a measure of a student's intrinsic valuation of their engagement with physics and enjoyment of this pursuit in a personally meaningful way [124]. Self-Efficacy (sometimes also referred to as competency belief) is a student's belief in their ability to succeed in a certain situation, task, or domain [18, 20], and may be associated with long-term student persistence [205]. The lower self-efficacy may partly be due to pervasive social and cultural stereotypes and biases and the paucity of positive encouragement and support endemic in the field of physics. All three of these factors – perceived recognition, interest, and self-efficacy – would, in general, contribute toward the development of a student's identity as a physics and science person [111, 124, 156].

There are several ways in which the development of a student's physics identity is important in the lab context. A student who develops a favorable and productive identity as a person who is good at physics is likely to engage, enjoy, and learn more in the lab [85], both during and after the course is finished. Low-stakes tinkering in the physics lab can be an important part of developing interest and self-efficacy in experimental physics and experimental science in general. A student's physics identity is valuable beyond the scope of the introductory physics sequence, even for students who pursue courses of study in which physics may not be directly relevant. For example, physics identity has been shown to be a strong predictor of interest and agency in engineering programs [111], and the movement toward competency-based assessments for medical school in some countries [258] makes clear that proficiency and confidence in using physics ideas and scientific ways of thinking are viewed as essential for future doctors.

A variety of prior studies have identified types of interpersonal interactions in labs and similar learning environments that impact student experiences differently according to their gender [62, 141, 180, 215, 237]. In one case, women in introductory college physics appreciated hands-on experiences as valuable but expressed frustration about having to adapt to new types of learning in the active engagement work employed in this class [180]. Research on gender in a robotics-based introductory engineering course shows differences in how women and men described experiencing certain learning activities and dealing with challenges, and also suggested that gender differences stem from the competitive aspects of the course [215]. In physics labs, observational studies have noted that women spend less time tinkering with apparatus [62, 141, 237].

We may understand why female and male students have different experiences in the same learning context by considering how and why a student may change their behavior in an attempt to align their self-image with societal preconceptions and cultural expectations of what it means to be a male or female physics student [115]. In other words, students will 'do gender' while 'doing physics' in order to conform to socio-cultural expectations. An identity framework is useful in this case because it provides a means to understand how identity trajectories of students who identify with different genders are shaped by their environment and experiences from the moment they enter the physics labs and how they position themselves and perform differently in the labs [120, 291]. Our goal, then, is to extend investigations that applied an identity framework to understanding how men and women worked differently in experimental physics research settings [61, 115] by applying this framework to the introductory physics lab setting. This can provide guidance for how to improve the introductory physics lab environment to make lab experiences effective for all students, at this crucial time when a positive boost in students' physics and science identity trajectories can set them on a path for growth as physics and science people and mitigate the impact of stereotypes that may otherwise thwart their positive physics and science identity development [61, 115].

2.3 Methodology

In order to investigate the introductory physics lab experiences and interactions of students who identify with different genders, we adopted a qualitative, mixed-methods approach that involved ethnographic classroom observations as well as semi-structured interviews with individual students. Both techniques are influenced by the reflexive strand of ethnographic investigation, in which the observer is mindful of their own positioning and background while planning data collection, interacting with participants, and analyzing results. It is through this reflection that blind spots, biases, and confounding preconceptions are identified and accounted for [39].

Both stages of this work were performed by the author and his supervisor, each with more than a decade of experience as a physics educator and a variety of personal experiences doing science in different cultures. The former is a graduate student, a White man and former high school physics teacher. The latter is an Asian female physics professor who has taught and conducted PER research since 1995. Throughout this work, these two investigators collaborated extensively to plan, conduct and analyze observations, and shared reflections at weekly meetings and frequently throughout the week as well.

2.3.1 Participants

The participants in this investigation are students enrolled in a stand-alone introductory physics lab at a large research university in the USA. The course is a one-semester introductory lab, which requires the second half of a two-semester introductory physics course as a co-requisite. Two versions of this lab, corresponding to the algebra- and calculus-based physics sequences, are offered. The algebra-based lab is often taken in the third or fourth year of study, and the majority of students who enroll are bio-science majors with an interest in health-related professions. Students in the calculus-based lab are typically engineering or physical science majors, and are more likely to be in their first or second year of study. While the algebra-based sequence is 55% female and 45% male, enrollment in the calculus-based sequence is 20% female and 80% male. University records at this time do not acknowledge non-binary gender identities.

The labs are run by graduate student teaching assistants (TAs), who are also responsible for most grading in this course. enrollment is capped at 24 students per lab session. Students are graded for completion of their work and, aside from a post-lab exercise, partners receive the same grade. The introductory physics labs have a reputation for being somewhat easier than other labs typically taken by students in this course such as organic chemistry, introductory biology, or introductory chemistry lab. Students who attend all 12 lab sessions typically receive at least a 'B' grade, and most receive an 'A' grade.

In both versions, students worked in groups of two (or three, if needed, e.g., if there is an odd number of students or some apparatus is broken so there are less stations available) to complete a thorough and detailed lab procedure during a 3-hour period. Our observations suggest that students self-select into partnerships essentially at random, as they sit down at an open lab bench on their first lab session. The exception is that a very small number of students partner-up before arriving in the lab: we generally see no significant differences in how these partnerships operate. Once formed, groups tend to stay together unless the TA requires a re-shuffling (see Section 2.6). Most students' pseudonyms were chosen by study participants: they reflect the participant's gender but not necessarily their racial or ethnic identity.

Table 1: Participants in the first study from the introductory physics lab, along with the pseudonyms of those quoted.

Gender	Female	13	(Leah, Elisa, Melanie, Bella, Natalie, Paulette, Zara, Liza, Janet, Kamala)
	Male	5	(Mark, Lou)
Major	Pre-Health Sciences	12	(Elisa, Melanie, Mark, Natalie, Zara, Liza, Janet, Kamala)
	Physical Sciences	5	(Leah, Lou, Paulette)
	Engineering	1	(Bella)
Course	Algebra-Based	12	(Mark, Elisa, Melanie, Natalie, Zara, Liza, Janet, Kamala)
	Calculus-Based	6	(Leah, Lou, Bella, Paulette)
Total		18	

2.3.2 Ethnographic Observations

The experiences that affect students' identity trajectories can be subtle and hard to identify. External observers, however, may be better-positioned to see how words, body positioning, and the manipulation of physical objects can contribute to student's experiences in the lab. We conducted observations many times over the course of the semester. These observations targeted six introductory lab sections during each of the fall 2018 and spring 2019 semesters. Each of the twelve sections was run by a different graduate student TA, who was informed in advance of the observation and asked to briefly introduce the observer at the start of the lab session. Observations lasted at least 1 hour each, in order to develop a fuller understanding of the student interactions that were being observed. In total, more than 100 hours of such observations were completed.

We took on the role of non-participant observers [228]. During our observation sessions, we sat on a side-bench of the laboratory and observed the students and TA while taking notes of what we saw and heard, as well as our reflections on what they might mean. An informal observation protocol [228] was adopted, and iteratively refined, as we sought to understand factors that might affect students' identity trajectories in the lab. With practice, and after comparison of notes between observers, we came to identify particular items of interest: comparing same-gender with mixed-gender groups, the work done by students in mixed-gender groups, and the nature of the students' discussions about their lab-work.

In line with our reflexive approach to investigation, we sought to fulfill three goals in how we positioned ourselves during our observations: acceptance, detachment, and reflexivity [39]. First, we aimed to position ourselves in such a way as to not influence the normal behavior of the TA or students. Sitting at the side of the lab helped in this effort, but we also engaged in a small amount of discussion with a few students (offering brief advice on the apparatus, asking Socratic questions about concepts, etc.) to establish the idea that we were friendly and unobtrusive. This was largely successful for the students, who were typically focused on their lab-work and ignored the observers. In follow-up discussions, some of the observed TAs agreed that our presence did not noticeably affect what the students did in the lab.

Our second goal was to keep sufficient distance between the lab participants and ourselves in order to make balanced observations. To this end, we kept discussion with students and TAs to a minimum (less than 10% of observation time). The third goal, reflexivity, required continuous reflection on how our own backgrounds and preconceptions may affect what we see, and what we deem important. To achieve this, we sought to pay attention to each individual student, to take their lab experiences at face value, and to discuss as observers these issues in order to come to a shared research agreement. 20 hours of our observations were done simultaneously, allowing us to compare notes and confirm that we observed similar events, behaviors, and interactions. In reviewing our detailed and thorough notes, we are confident that we were successful at maintaining suitable detachment and reflexivity in our observations.

2.3.3 Semi-Structured Interviews

Based on our classroom observations, we identified students whose perspectives and experiences we thought would (a) provide a cross-section of the students who enroll in the lab classes, and (b) had experiences and perspectives that would be valuable for us in understanding student interactions in the lab. These students were invited to participate in hour-long interviews, for which they were compensated with a $25 payment card. Roughly half of the students who were invited agreed to be interviewed and we conducted a total of 18 interviews at the end of the fall and spring semesters during the 2018/19 academic year.

Our reflexive ethnographic observations suggested differential gender effects with negative impacts on women, so we aimed for an interview pool that included more women's voices. Our decision was supported by the fact that only two of the five men we interviewed were aware of these effects (perhaps experiencing a blindspot [17]) while all of the women were able to describe at least one way in which men and women experienced the lab differently. In addition, we sought particularly to speak with students from mixed-gender partnerships, as these seemed to be the locus of gendered inequity of opportunities, based on the ethnographic observations. By comparison, we observed that students who worked in same-gender groups tended to collaborate much more effectively and equitably. Of the 18 participants who agreed to participate in interviews, 13 identified as female and 5 as male, and all but one described working in a mixed-gender group for at least part of the lab course (we note that most students in the lab course stayed with the same partner throughout and only a few occasionally switched). All 18 participants worked in groups that were stable over the course of the 14-week semester.

Drawing on our observations we assembled and refined a list of potential interview questions to serve as our interview protocol [228]. These included questions about the student's background and prior lab experiences; interactions with other students and the TA; thoughts on the structure, mechanisms, and effectiveness of the course; and experiences with task division, including gendered division of labor. Despite the long list of questions, we sought to make these semi-structured interviews conversational in nature to give participants the opportunity to express themselves freely, dig deeply on critical issues, and remain comfortable and safe. The investigators used the list of questions to gently steer the conversation in the directions specified by the interview protocol. Most participants required little prompting and were keen to share openly. All interviews were audio-recorded and transcribed.

2.4 Results and Discussion

2.4.1 The Secretary and the Tinkerer

Students in the physics lab have a wide variety of background experiences. Some have taken AP Physics in high school, while others went to schools that didn't offer it. Leah, a high-achieving chemistry major, described why she didn't take physics in high school:

> I had never had physics in high school at all. My school pushed for biology and chemistry for girls, and physics for guys... So when I came here I had no clue about anything about physics... I was clueless in a sense about physics. Physics I and II, the calculus-based ones, were [a] little fast for me but a good speed for everyone [else]. The physics lab seemed a lot slower paced, so it was really good for me but it was kind of boring for other people that were very, very, very good at physics...

Leah's high school experiences established a clear picture of who can be a physics person, so it is unsurprising that she expressed a low level of physics self-efficacy and did not see herself as a person who can be good at physics. Furthermore, her prior preparation meant that when she got to college, Leah had little confidence in her ability to do physics. Her low self-efficacy is clear when she compared herself with peers, whom she perceived to be mastering physics concepts much more quickly than she was. However, Leah acknowledged that when she compared her grades with those of her more-confident classmates, she saw that she was doing just as well as them.

> There would be times when I would feel like I am not good at physics, I am not good at it. But we would get tests back... I was very comparable to them, but I still felt like, 'Oh, it's not my thing, I'm not very good at it.' But here I am, and they think they are very good at it and I'm doing just as well as them.

While Leah was certainly doing well in class, her physics identity was stagnant because her low self-efficacy prevented her from internalizing the idea that she was developing mastery of physics concepts. Even though Leah was telling her (male) peers that they were "very good at physics", no-one was communicating that type of message to her or recognizing her success. This conflict is typical for women enrolled in introductory physics and even though men and women perform equally in introductory physics at the institution where this study was carried out, men report substantially higher self-efficacy [305]. In negotiating a role in the physics lab, it is Leah's low self-efficacy as a physicist – developed through the lack of support at high school, the encouragement she didn't receive as a student, and a shortfall of recognition when she did begin to demonstrate mastery – that may have led her to adopt a secretarial role.

Mark, a microbiology major in his final semester, also hadn't taken physics in high school. However, unlike Leah, Mark was given opportunities to play with circuits as a child, and to learn how to work with electronics through school research programs. These prior experiences helped reinforce Mark's interest and self-efficacy in science.

> I've taken apart a lot of things. I've done work with Arduinos, kind of, building my

> own circuits... [My father is] a chemical engineer, so I always had something I could work with when I would take things apart, until I bought my own things... some of [the Arduino work] I had done with my research experience outside of school, having to design some things, measuring bacteria and things like that. One of them, I did this summer program where we build a little thing to switch LEDs off and on, and also to measure absorbance inside cultures.

Mark's prior experience led him to adopt the role of the Tinkerer in his lab group. He recognized that this meant an unequal division of labor in his group. When asked explicitly about why male students sometimes took over the apparatus in the introductory lab, and what could be done about it, Mark replied:

> I would say maybe some of the labs that had a more technical set-up, I would do more of that. And then while I was setting that up, she would be waiting... I'm usually in front of the machine so I'm usually handling that while she's inputting all the data. And that's maybe something to think about, maybe changing the roles.

We see here an example of masculine lab behavior being replicated along gender lines [115], to the benefit of Mark at the expense of his partner, Elisa. Furthermore, Mark attempted to blame his dominance of the apparatus on his seating location. We found this attribution by male students when questioned to be common, but spurious, as we observed most students alternate locations readily as they do their lab work.

Elisa agreed with Mark's description of the unequal task division in their partnership, but speculated he must have taken advanced physics classes to have such a high self-efficacy with the lab apparatus (in fact, he had neither taken high school physics, nor had he taken any physics classes she had not). Here, again, notice how asymmetric engagement with the lab-work only provided opportunities to Mark, potentially bolstering his physics identity development while hindering Elisa's. By assuming that he must have taken advanced classes and allowing him to do the tinkering, Elisa appears to recognize Mark's practical skills and self-confidence. This is a message that may have bolstered his self-efficacy, even though he hadn't actually taken such classes.

On the other hand, Elisa was doing the other work: she didn't get to develop expertise with experimental techniques in a low-stakes environment, didn't get acknowledged by her partner positively, and therefore didn't get an opportunity to develop her physics and science identities. The types of task they each performed and the opportunities she and her partner had in the physics lab are likely to further increase the gap between their self efficacies when it comes to tinkering and the associated learning in the lab. Elisa elaborated, describing a typical day in the lab as follows:

> He liked to do a lot of the setting-up and he knew what was going on, more than I did. I felt like we both tried to split [it] up, so it wasn't one person doing all the work. I like to do the data entry and stuff, so often I would do that.

This division of work into Tinkerer and Secretary roles was a theme we saw repeated frequently, in both algebra- and calculus-based labs, when students worked in mixed-gender groups. In most cases, the Tinkerer tended to be male, and the Secretary tended to be female. When the Secretary-Tinkerer split happened, as with Elisa and Mark, students typically thought of it as a fair division of labor. Melanie, a biology major, described how she and her partner split the work:

> In lab, I'm usually in charge of writing down the data that we collect, and he's usually the one doing the physical part.

While the Secretary-Tinkerer task division looks fair on its surface, there are two big reasons why it can be a deleterious approach to work. First, this division can reinforce a power imbalance in team-work that deprives the Secretary of the opportunity to be a scientific investigator. Lou, the partner of Leah (above) and a fellow chemistry major, described a moment when he interfered with his partner's attempts to contribute to building a complex circuit:

> Sometimes I get a little carried away with getting things to work. If Leah would come over and try to change things, I'd be like, "I've almost got it." That's just my personality.

Leah described the same type of interaction in her interview. Traditional gender roles were being enacted here: the man as authoritative, and the woman as responsive. However, Leah wanted to do her fair share of the tinkering and recognized the inequality in their division of the work. The following situation was a rare case of the Secretary being willing to speak up and risk conflict, and may be seen as arising from a mismatch between Leah's relatively high level of initiative as a learner and the expectation that Secretaries have a more passive role.

> In the circuits labs, he kind of took over the experiment... the next week, I was kind of like, "okay, give me that wire." I tried to do more of the trying to plug in and see what's going on.

The second reason the Secretary-Tinkerer split is deleterious is that it deprives both members of practice with the other type of working. Since the physics lab is often the only place students learn to do hands-on experimental physics, the Secretary stands to lose more from this task division than does the Tinkerer. Many of the skills recommended by the AAPT [1] such as constructing and using apparatus, making measurements, and troubleshooting problems cannot be learned by watching a partner. As a contrast to her introductory physics lab, Bella described a digital circuits lab she took as an engineering student:

> It's mainly the guys who are building the labs. And the women are mainly having to figure out the software and the calculations... I don't know, maybe it's a perception that men are better at things that require the use of hands?

When asked whether the gender split deprived women of opportunities to learn, Bella explained that she felt under-prepared for a mid-term practical assessment in her engineering lab:

> Definitely! It definitely does. On the practicum, I remember thinking, 'Dang, my partner always did this part of the lab.'

Although most of the interview participants discussed short-term impacts, the Secretary-Tinkerer split can also have long-term negative consequences. In particular, this inequitable task division deprives women of the opportunity to tinker in a low-stakes environment, which is necessary for developing one's physics and science identity as a person who can handle the equipment and experiment.

2.4.2 Hermione and the Slacker

While the Secretary-Tinkerer mode of work deprives female students of the opportunity to tinker with apparatus, which is a critical part of the lab and essential for identity development as a physics or science person, we observe a very different effect in a second mode of work that is equally salient. In this case, a student, typically female, ends up shouldering a disproportionate amount of the work and compensating for the shortcomings of their partner(s). Such students take on the responsibility of ensuring the work gets done when their partners fall short, but are more than just a project manager. In the physics lab, Bella described working with two partners and asking one of them a question, only for him to turn the question back on her because he hadn't prepared for the lab and did not want to think about it.

> I feel like I did a lot of the thinking for the group... [When I asked him a question] he would be like, what do you think?

Typically, students who adopt this Hermione archetype see it as necessary in order to complete their lab-work because their partner, the Slacker, appears to be uninterested. Like the Secretary-Tinkerer split, the Hermione-Slacker task division is one that seems to strengthen as partners work together for more than one lab session, as the partners recognize that the other person would be willing to pick up the slack.

The Hermione-Slacker mode of work seems to be especially prevalent in groups of three students, although we also observe it in pairs. It may partly be that the student(s) realize that their lab partner will make sure things get done and, thus, they will receive a good grade with minimal effort. Natalie explained her disappointment that her partner wasn't contributing:

> I like being on a team... Seeing that he puts in as much effort as I put in... Because I don't see that effort coming from him, I've had to step up to make up for that effort so we get it done with.

The lack of engagement or initiative from Natalie's partner, however, went beyond merely not contributing. She described how her partner's disinclination to participate led to her skipping a portion of the lab report that was not explicitly graded:

> In the beginning of the semester, I would try to do the analysis questions just because I wanted to understand it more, and he was like, we don't have to do this, there's no reason to doing this. So I kind of gave up on that portion.

As a result of this partnership, Natalie's opportunity to grow her expertise and interest in physics was stymied, and in the rest of the interview it was evident that subscribing to her partner's lackadaisical approach to doing the lab just to get a grade may have negatively impacted her physics and science identity development.

Despite being a physics major, Paulette's male partner seemed to have little interest in completing the lab, let alone contributing equally to the mental and physical labor required to complete the work. This put her in the awkward position of needing to repeatedly ask him to contribute to work for which he was receiving a grade and, perhaps worse, forced Paulette into a traditional – almost maternal – role, depriving her of the opportunity to dig deeper and develop her self-efficacy as a subject-matter expert.

> Well, my partner's a little lazy... Sometimes he's on his phone and stuff, and I'm just like, 'get off your phone.' He helps when I ask. I'll be like, 'hey, can you do this?' But he doesn't really start doing stuff himself most of the time. I'm like, 'I'm not your mom.'

As time went on, Paulette explained, he took increasingly-long and increasingly-frequent breaks from the lab, and contributed less and less to the lab-work they should have been sharing equally. She described asking him to help, but he was so detached from the entire task that he would not even know where they were in the lab procedure or what needed to be done.

We observed Hermiones taking on a variety of tasks, including preparing for lab when their partners did not, managing the work-flow, assigning small tasks to their partners and monitoring their progress, communicating with the lab TA and other groups, and ensuring the data collection was complete before leaving the lab room. We also saw Hermiones take on the labor of reconciling different and sometimes conflicting instructions, methods, and conceptual ideas. It added up to a lot of commitment and effort, and so frustration with a partner's lack of preparation is a common theme for students such as Zara in this role. Here, she described what it was like when her partners didn't adequately prepare for the lab, and her experience the one week the group had to stay late in order to finish their work because she wasn't as well prepared.

> There was one lab where, working with circuits... that was very difficult for me. Maybe it's just because during the week I didn't have a very good week or something. I really struggled understanding it. I think my partners weren't always prepared for the labs, so it fell on me to understand and get the group to finish the lab... I need to be prepared to know what's going on, because they won't.

Despite the disproportional amount of time and effort she invested into the lab, Zara either didn't receive or didn't internalize recognition from her peers. When asked if she was the expert in her group, she laughed and said:

> I definitely would not call myself an expert. Maybe I read the lab manual more?

According to the identity framework, perceived recognition should stimulate development of Zara's physics identity. However, because her lab participation was managerial, rather than focused on the physics or hands-on parts of the lab work, the recognition she received from her partner was – in her view – related to the project management, rather than mastery of physics concepts and skills. Moreover, it appeared that Zara wasn't internalizing the little recognition she did receive from her peers, and so she appears to have experienced little identity development as a physics person.

Like Zara, Liza described her Hermione role in a way that situated her as doing necessary work to accommodate an unprepared peer:

> He didn't read the manual every week, a lot of the time it was me telling him what to do... do this, do this, and it would be me doing the note-taking... I felt like I was controlling from that position.

The Hermione archetype can disadvantage students who adopt it in part because they do the majority of the work while receiving the same learning experience and/or grade. Even worse, the managerial work they do takes them away from the tinkering and sense-making activities that could help them to develop their identities as physics and science people. Janet described spending a large portion of her time mediating between her partner and the TA, asking questions to the TA about things she already understood, in order to appease her partner after he hadn't bothered with the pre-lab reading and expressed doubts about her explanations of the tasks they needed to do.

> It's like, you're wasting my time because you're unprepared. Well, now I'm not able to learn as well because I'm spending so much time asking [his] questions [to the TA] that I don't really need to ask, because I know what's going on. It's wasting my time...

Since Hermione-role students are typically situated as the hard-working one in their partnerships, these students tend to attribute their successes to their exertions rather than their physics competence, which could again shortchange their physics and science self-efficacy and identity development. And because they need to be so laser-focused on getting everything done for their group, there is little time or capacity for Hermiones to develop higher levels of self-efficacy and interest in physics, and to grow their physics and science identities, through their experimental work.

Kamala, a high-achieving pre-med student who managed a group of three, praised the skills of one of her partners:

> He's very good at equipment, so even if he doesn't necessarily read the lab, he's just one of those people that has very good problem-solving skills when there's hands-on things.

On the other hand, when it came to her own expertise, she rebuffed credit from her partners, interpreting what they say as not genuine, saying:

> They have an impression that I'm just better at physics than they are. Or I'm just smarter at this stuff than they are. Which isn't necessarily true. It just comes down to... are you willing to push the group forward in terms of knowing what the next thing to do is?

In effect, then, Kamala praised her partner for practical work, which he did because of his confidence with the equipment but without reading the lab manual, while she appears to have internalized no recognition for her mastery of the physics concepts or experimental procedures. In part, this was because she felt she was essentially managing the lab work for her group in order to make sure it got done. This is a common theme in these interviews: women displayed lower self-efficacy than men, and were more likely to attribute their success to external factors such as hard work rather than to their own developing mastery of experimental physics. By focusing on managerial work, women who adopted the Hermione archetype received recognition that was either not relevant to their physics and science identities or that was interpreted as not being genuine. They consequently appear to have experienced physics and science identity growth that was stunted in comparison with their peers in same-gender groups, or in comparison with the men in the class who adopted Tinkerer roles.

2.5 General Discussion

While the Secretary-Tinkerer mode of work has been documented before in research on STEM education [24, 41, 174], here we introduce the Hermione-Slacker mode for the first time. We believe that this taxonomy will help educators conceptualize and reflect on the ways in which students may be disadvantaged by gendered modes of work in the physics lab and other places in which students are doing science together. These archetypes are both salient and ubiquitous in mixed-gender groups, especially when compared with same-gender groups.

Applying the identity framework, we find that both Secretary and Hermione archetypes can act to stunt the development of physics and science identity for women in these roles. Women in secretarial roles, like Leah, Elisa, and Melanie, are denied the opportunity to actively engage with the apparatus in a low-stakes environment of the lab, and thus do not benefit from this opportunity to grow their interest in experimental science. Thus, it is unsurprising that Secretaries typically describe a transactional view of their lab-work: they do what is required, and do not see themselves as undergoing growth in their identity as physics or science science people as a result of the lab course.

In the same way, discussions with women in the Hermione role, like Natalie, Paulette, Zara, and Kamala, suggest little growth in their physics and science identities as a result of this physics lab. Because they are pushed to adopt a managerial (or even maternal) role, they see themselves primarily involved in getting things done, leaving little time to deeply engage with work that might stimulate growth in their interest and self-efficacy in physics, or the consequent development of their identities as physics and science people. And while their partners sometimes recognize them for their leadership in the introductory physics lab, they rarely appear to internalize those types of recognition for their accomplishments in terms of being good at physics. In total, of approximately 20 students in the Hermione role we identified during the observation phase of our work, none of them were men.

Overall, then, we find that students whose negotiation of “doing gender” and “doing physics” results in them adopting Secretary and Hermione roles experience interactions in the lab that limit their physics and science identity development. Returning to the identity framework, we note that students who adopt the Hermione or Secretary roles receive inadequate recognition as scientists. Hermiones, in particular, may receive recognition from peers that is either not perceived as genuine or inadequate compared to the work they do. Likewise, the task division encountered by students in both these roles may have provided fewer opportunities to develop an interest in experimental science, but this was less explicit and salient in the interviews. Finally, students in the Hermione or Secretary roles spend time on managerial or notekeeping work that doesn't promote development of their self-efficacy as scientists. Since these two roles tend to be occupied primarily by women, given the existence of pervasive societal stereotypes about physics that can disadvantage women, this issue deserves careful attention.

We emphasize that this analysis is focused on the impacts of gendered roles in the lab on students. For example, many students in Tinkerer-Secretary partnership may have good intentions. In particular, some students who adopt a Tinkerer role may view themselves as doing extra work to the benefit of their partner, while their Secretary partners may believe that stepping back from the apparatus allows them both to finish the lab efficiently.

Some of the classroom observations that the author and his supervisor conducted were at the beginning of the semester (first lab class) to observe how students selected their partners and settled into different roles. Based upon these observations for mixed-gender and same-gender groups, we propose a model for how the gendered-roles solidified in many mixed-gender groups compared to the same-gender groups.

We present in Fig. 1 one possible way to visualize the dynamics we identified during our observations that were corroborated by interviews. In this model, a student's initiative - their willingness to do work - in the lab is plotted horizontally, while the vertical axis shows a student's gender. Fig. 1a shows the typical dynamics we observe when a woman with lower initiative begins to work with a high-initiative man: he tends increasingly to take over the experiment, adopting the Tinkerer role, and she tends more toward the Secretary role. Similarly, Fig. 1b shows what we typically observe when a high-initiative woman begins to work with a low-initiative man: she adopts a Hermione role, and he becomes a Slacker. We observe this type of dynamics that drives this task division throughout the lab period, but they are especially pronounced during the first hour that a pair of students is beginning to work together.

Our observations suggest that unlike in mixed-gender groups, the symmetry breaking and "phase separation" into different roles generally does not seem to occur in same-gender groups. In fact, in our observations, the general contrast between the mixed-gender and

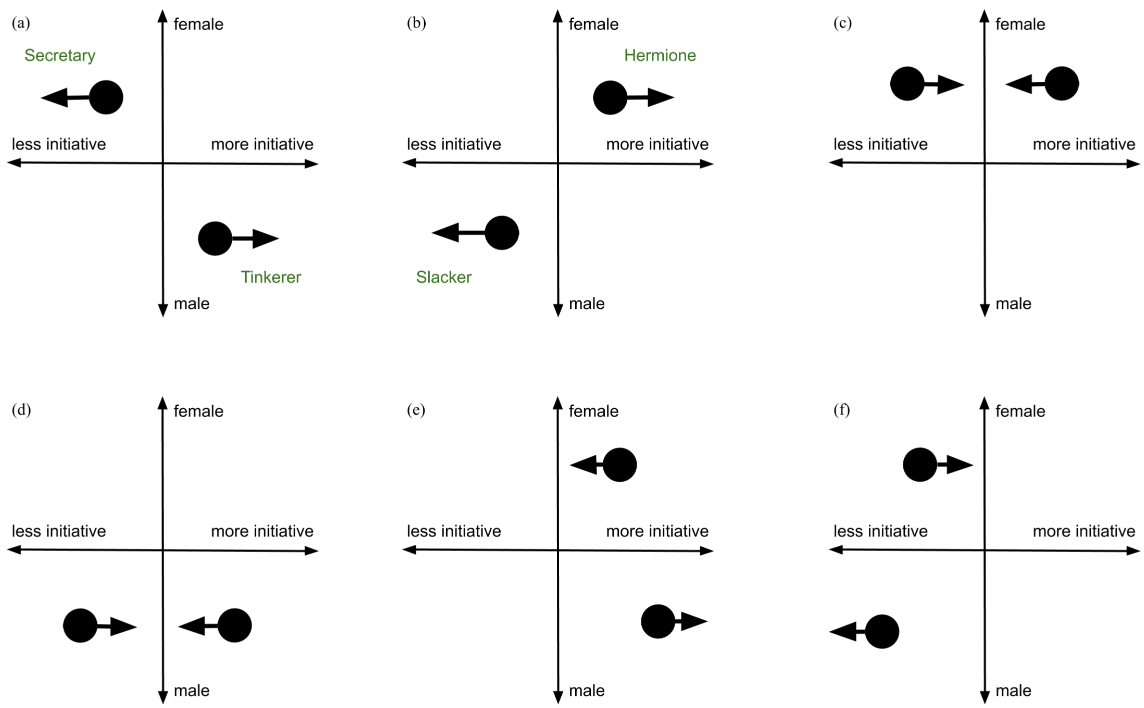

Figure 1: A proposed model to account for how female and male students settle into adopted gendered modes of work in mixed-gender groups.

same-gender groups in this regard was striking. As depicted in Fig. 1c and Fig. 1d, typically, two students of the same gender who work together - regardless of their initial differences in initiative - tend to achieve an equilibrium, adopting similar types and amounts of work. In these same-gender partnerships, there is little psychological distance [292] (a measure of the similarity between two people based on their characteristics, their behaviors, and the social groups to which they identify) between the partners. This may be a relevant factor in determining whether two students will collaborate effectively.

We also noticed, in our observations, a few cases of mixed-gender groups that began the lab period with comparable initiative. This was the case for Leah and Lou, as described in the previous section. Even though they both started off with a high level of initiative, the female student appeared to display slightly less initiative (which may be due to her gender identification in the mixed gender group), and this small difference was exacerbated by the collaboration, as shown in Fig. 1e. However, Leah's determination to take an active role in the lab-work meant that neither she nor Lou moved very far on the diagram, but also introduced tension to their interactions sometimes as discussed earlier. We find that when there was tension in groups, it typically came from conflict between students' desired form of participation in the lab and the role-division described here. In another group, we observed two lower-initiative students (compared to the average of the class) struggling to complete the lab-work until the student with slightly-more initiative started to put in more effort. In this case, shown in Fig. 1f, the female student appeared to have slightly more initiative originally, and was pushed toward adopting a Hermione role while her partner became more of a Slacker.

In summary, we propose a preliminary model of lab dynamics in which gender identification of students, as a type of psychological distance, acts to push similar (e.g., same gender) students toward fair and equitable work-patterns, while driving dissimilar (e.g., different gender) students toward the inequitable archetypes described above. Future work is needed in order to refine this proposed model and understand how labs in which students partner with each other may produce the type of dynamics we observed. This may include looking at constraints and affordances of the lab such as how the grades are assigned and the impact on students' physics and science identities of emphasizing the long term utility and value of dividing all aspects of the lab equitably. Furthermore, while this chapter has emphasized gendered instances of these archetypes, we note that task division in our labs can also be influenced by our students' racial and ethnic identities. Unpacking and understanding this effect would require an intersectional lens, and is beyond the scope of this chapter.

2.6 Implications for Practice

A key question is how to address these inequitable modes of task division. Below, we describe five approaches we have started to implement in our labs, which appear to be promising. In our observations, these approaches seem to be beneficial for students from all four archetypes identified above. Just like Secretaries and Hermiones, Tinkerers and Slackers benefit from increased accountability (including grade incentives), more clearly defined responsibilities, and opportunities to renegotiate their role in group work.

First, regularly changing group composition may help to reduce some types of inequities in group-work [125]. When students work together over several weeks, we see that their adoption of inequitable modes of work (including task division) becomes solidified over time. In labs that changed groups mid-semester, our observations suggested more-equitable work in the second half of the semester compared with the first half of the semester.

A second, often-recommended, approach is to assign (and rotate) roles within student groups [126]. Recently, we have observed some success in reducing the occurrence of some inequitable task division in labs where the TA (after being prepared to do so via professional development) required that student partnerships take on the roles of 'experimenter' and 'recorder', and alternate weekly. Students were generally willing to play along and stick to their roles, but we noticed that some recorders would take over parts of the experimentation role if their partners struggled. Likewise, students who were recorders during a given lab session had no opportunity to develop practical skills with the apparatus for that particular lab, and so even though they might get as much experience with the apparatus as their partner, neither of them gets as much experience as they would both get if they collaborated equitably. A series of 'checkpoints' provided a grade incentive to students to ensure they fulfilled their roles.

Third, we found that isolated minorities – such as a woman in a group of three with two men – were particularly vulnerable to the archetypes described above. Thus, we endorse the advice [126] to avoid isolating minorities if possible. However, this idea can be usefully extended by considering what happens when a minority student is in a class without any peers from their minority group. During our interviews, one student described her experience as the only female student in an honors physics lab. Despite being friends with many of her male peers, none were willing to partner with her for investigations that were too comprehensive to effectively complete alone. She eventually dropped out of that course and enrolled in the regular (non-honors) lab the next semester, but told us that she would have stuck with it if she had had a fellow female student with whom she could have worked. Thus, we suggest that instructors be careful not only to avoid isolating underrepresented minorities in groups of non-minority students, but also to take care not to allow underrepresented minority students to be isolated without the social resources they need to complete their work with the same level of support as other students.

Fourth, we note that a small amount of previous experience can provide a big boost to a student's self-efficacy when it comes to lab-work at this level. Mark and Lou attributed their tinkering predispositions to extracurricular science activities they experienced at school, for example. One approach, then, is to ensure that all students have an opportunity to tinker unimpeded during the first few minutes of the lab session. Since women are less likely to have had such experiences due to societal biases and stereotypes [59], perhaps the lab room could be equipped with enough apparatus for each student to build a few simple circuits, use a caliper, or set up a lens individually before undertaking the cooperative part of the experimental work. In this way, individual 'tinkering time' could be built into the lab-work, and students could be coached to do it alone, and not to interrupt their partners' preliminary tinkering.

Finally, recognizing that collaboration is a skill like any other, we have begun to develop and systematically evaluate lab tasks that explicitly divide the learning tasks between partners. For example, Student A is assigned to develop the theoretical prediction while Student B carries out the measurement before they share, and switch roles for the next part of the experiment. This structured work can act as a scaffold over the first few weeks of the semester, and can then be slowly withdrawn as students become more familiar with the expectations for equitable collaboration in the physics lab and more capable of working in an equitable way over the course of the semester. Here, too, we adopted a mixed grading scheme that partially accounted for individual contributions to a group lab report.

As a further issue, introductory physics labs at large research universities are often run by graduate teaching assistants. In such cases, professional development to establish 'buy-in' [112, 316] for the principle of equitable learning, designing approaches for teaching assistants to use in their labs, and instructing and monitoring the use of these approaches will be an essential consideration for such settings [73]. In labs run by our graduate student TAs, we see little or no impact from any of the above strategies when the TAs do not believe in their necessity and benefit.

While the adoption of the archetypes described and illustrated in this chapter – the Tinkerer, the Secretary, the Slacker, and Hermione – is symptomatic of inequitable learning in the physics lab, these archetypes also serve as a way to understand the nature of the inequities. It is our hope that these labels provide a vocabulary for discussing equity in the lab and a reference point as we work to transform introductory labs into places where all students develop positive identity as physics and science people.

3.0 Experiences of Women in Physics

3.1 Introduction

Issues related to why women are underrepresented in physics are being investigated from various angles (e.g., see Refs. [44, 124, 245, 246, 254, 319]). While some women experience explicit and implicit bias while studying physics, some may experience being recognized as a person who can excel in physics and may experience mentoring by physics instructors [148, 201, 205, 225, 253] differently from men even in the absence of discrimination [151]. Women can also be affected by isolation and stereotype threat [156, 195, 197]. Here we present the voices of two white American undergraduate women in their second and third years at a large public university to shed light on experiences in physics that they found alienating. Both were interviewed about their experiences for one hour each using a semi-structured protocol. All names are pseudonyms.

3.2 Leia: Self-Efficacy and Recognition

Self-efficacy, the belief in one's ability to be successful at a given task [19], has been shown to be an important factor in students' career decisions, enrollment in STEM courses, and persistence toward long-term goals [225, 253]. However, research has shown a sizable gender gap in the self-efficacy of students enrolled in introductory physics courses, in which women consistently have a low self-efficacy than their male peers, even if those men have lower grades [205]. Leia explains what this paradoxical self-efficacy difference looks like, from a student's perspective: "There would be times when I would feel like I am not good at physics, I am not good at it. But we would get tests back... I was very comparable to them but I still felt like, oh, it is not my thing, I am not very good at it". While Leia knew, on some level, that she was doing just as well as other students in her class, at another point in the interview she described hypothetical physics students who were far superior to her:

"Physics I and II [lectures] were a little fast for me but a good speed for everyone... but the physics lab seemed a lot slower-paced so it was really good for me, but it was kind of boring for other people that were very, very, very good at physics."

While Leia had a low physics self-efficacy, she was confident about her capabilities in chemistry. She recounted how getting a very good grade on the physical chemistry (P-chem) midterm exam in the current semester was something that boosted her self-efficacy and increased her interest in P-chem to the point where she was almost looking forward to her final exam: "I just got the grade back. I almost cried because this is the best I have done on a chemistry test. And then I just felt so much more confident to take the final. You know, like, I am REALLY good at P-chem... It was really nice, like, YES! All my hard work is paying off finally." Leia clearly has the ability to be successful, and she has a mindset that supports her in this closely-related field to physics. However, the fact that she feels so negative about her physics abilities serves to highlight how threatening an experience the introductory physics class can be for some women, even though all students are ostensibly treated equally.

In order to understand how Leia could have a positive view of physical chemistry while being apprehensive about physics, we asked her about the recognition she has received as a chemistry student. Students who perceive that they are being positively recognized as members of a science community often experience a boost to their identity as scientists [156, 190]. Leia noted that in her introductory physics course she found support from a female undergraduate teaching assistant from the Undergraduates Teaching Undergraduates (UTU) program. She felt that this was a person who was empathetic, and who understood her and her questions. She also felt she could talk to the teaching assistant about physics at her level and she did not feel anxious talking to her. She was so positively impacted by the support she received from the female physics UTU experience as a student that she herself decided to be a UTU in physical chemistry and noted that the recognition from the students she was teaching was the best part of the experience. "I think honestly the best part was the last session. People were excited to leave, it was the last lab, but then they would turn around and [say] 'oh but we would never see you again' ... that was the best part. I think I made their experience in the lab like a little bit better." Further discussions suggest that these UTU experiences, both from the perspective of a student (asking a female UTU for help in the physics context) and as a UTU teaching students in the chemistry context, were empowering for her, and that the recognition she received from both types of experiences helped improve her sense of belonging and self-efficacy.

When discussing what we can do to help students like her, Leia talked about the small class size in her physical chemistry course this current term, "My class size got smaller so I felt like the professor knew my name. That was the first time, well in a science class, a professor really knew my name... So it is just that the higher you go, the smaller the classes, the better you get to know the people. That is the best part for me." Leia further added that smaller classes would really help her because the instructor will be able to communicate with her personally saying, "Let's say I did really poorly 'Leia, what happened' or if I did really well 'Great job, Leia!' ... I like that a whole lot better than huge, huge classes...". She then reminisced about a non-science course saying, "I had a public speaking course which is my favorite course here so far and [the instructor] is phenomenal! She would cross the street to just say hello. So yeah it definitely, definitely made me feel really welcome." It appears that a smaller class helped to improve Leia's sense-of-belonging and self-efficacy by making it more likely that she would experience positive interactions and receive recognition as a scientist from her instructors. This recognition, once internalized, may have reduced the ambiguity she felt in how she was performing relative to others, and may have helped increase her self-efficacy and sense of belonging. This process may be especially relevant for some students who may not have previously had opportunities to learn physics and interact with content area experts, such as students like Leia who attended small or under-resourced high schools.

3.3 Paulette: Isolation, Stereotype Threat, and Mentorship

Another difficulty faced by women in physics is the challenge of isolation. Being the only member of a visible minority in a cohort, lab, or department can lead a student to being excluded from productive study groups, missing out on informal forms of support, and even having vital materials withheld by peers [246]. Paulette's experiences in the introductory honors lab are an example of this. She was the only woman in the lab course and consistently found herself working alone, even though she was friends with men in the lab. "[The lab] was really, really cool but it's very unguided... I went in a month and was like, I can't do this anymore. I was the only girl in that class, which is… not ideal. It just doesn't make you feel very comfortable, ever." Paulette explained how she ended up working alone by describing how the male students paired up amongst themselves: "[The other students] were like, hey do you want to do the photoelectric effect with me, do you want to do blackbody radiation with me? … I already knew them, and that was still kind of hard." As the semester went on, Paulette's experiences in the lab became frustrating and exhausting: "it's like, you say stuff, and people don't really listen to you. It's not necessarily intentional. But you suggest something, and people just ignore you. So, you have to be a bit more forceful, and sometimes you just don't have the energy for that… after a certain point I didn't want to go to that class anymore, and it was, I'm just going to withdraw. I didn't want to deal with the mental stress of that. Or, you know, pay money for a class I didn't want to go to".

Paulette also faced the challenge of stereotype threat as a woman in a male-dominated field. Stereotype threat refers to fear of confirming a negative stereotype about oneself because of one's association with a stereotyped group (e.g., women in physics). Stereotype threat can increase anxiety and rob students of the limited cognitive resources and thus lead to deteriorated performance. Stereotype threat has been posited as a possible explanation for the underperformance of traditionally marginalized groups in physics [195, 197, 284, 285] and Paulette's experiences provide a vivid example. Paulette began by articulating a stereotype threat that impacted her success in class: fear that her instructors were judging her for being a woman and being condescending in their responses to her questions. She mentioned that some of her male physics professors "tend to have an air where it's like, this is the naïve solution, this is so trivial... Sometimes there's condescension, which makes you not want to ask questions anymore." Consequently, she feared raising her hand to ask questions in class: "It can be especially hard when you have that question in the back of your mind which is like, are they being condescending because they think I can't do it, or because I'm a woman?" This fear and response became a threat and the condescension had a negative impact on

Paulette's sense of belonging. She explained how she, and her female classmates, experienced a lower sense of belonging than their male peers, and how this impacted them differently: "I feel like most of the physics majors at one point are like, I'm too dumb, everyone else is smarter than me. But I've only ever heard the women contemplating… I don't belong in this and maybe I should drop the major. I've never heard any of the guys say that. And I don't know… there's only one thing that can really stem from!" Paulette described this gender difference as arising from feelings of isolation, gendered microaggressions, and the ubiquity of stereotypes such as the belief that physics is a field for 'brilliant' men. We can imagine how the same process may have hurt Paulette in other circumstances by noting that, since human working memory has a limited capacity while problem solving [249], lower self-efficacy and the resulting anxiety associated with stereotypes could have robbed Leia of some of her cognitive resources while learning physics or taking physics quizzes and exams, and she could potentially have done even better than she actually did if she did not have such challenges.

Paulette also described positive experiences during our interview, such as how she feels empowered with the support of a female faculty mentor. Mentoring can play an essential role for students, providing essential relational resources to support the development of students' sense of belonging and identity as scientists [29, 148]. Similarly, role models can help to counter stereotypes about who can succeed in physics. In her interview, Paulette noted that both faculty mentors and peers can serve to provide support and reassurance. Paulette lamented not having a female professor in her early physics classes: "I think it would be nice to have a female professor... because I [would] just feel more comfortable talking to her. Until then I've just had men." When asked what kinds of things have supported her in college so far, Paulette described her current research internship with a female professor, "I have my female mentorship thing with [a professor of physical chemistry], which was definitely on purpose. I looked for female PIs." She was very happy about her relationship with her mentor and noted, "We meet every week. She's like, how are you doing, where are you at? … The times I know I have problems, I'll talk to her then, she'll talk me through it, and I can keep doing my work." Support can also come from peers. Reminiscing on the lab Paulette said, "I know, almost for a fact, that if my friend Amy had been in that class I would have stayed in that class because we would have partnered on pretty much everything." Paulette also found comfort in talking to a female physics teaching assistant, "My first semester, the TA, she's a graduate student here. That was nice, I just automatically felt more comfortable asking questions."

3.4 Implications and Summary

Reflections on the interviews with Leia and Paulette lead to the following suggestions for instructors:

1. Ensure that students are not isolated. In settings where group work is expected, or necessary, ensure that students partner equitably even if it means being "uncool". As Paulette suggests, "making the teachers assign partners, as uncool as that is, would definitely help with that kind of thing. Because then you don't feel like you're forcing someone to be your partner." Just as research suggests that instructors be careful to avoid isolating underrepresented minority students in majority-dominated groups [126], we suggest that instructors should also be careful to avoid isolating underrepresented minority students into doing solo work because they cannot always find a partner.
2. Make sure to recognize students for their successes, including the often-overlooked success of simply making progress in their studies. Also, take interest in students' well-being. As Leia describes, simply acknowledging students inside or outside of class can have a long-lasting positive impact. For instructors, this could mean learning students' names and using them, acknowledging students when you see them outside of the classroom, and providing friendly encouragement when students do well on an assignment.
3. Take responsibility for the social dynamics in your classroom and pay attention to students who may be feeling under-supported. As Paulette explains, "I think it was probably pretty obvious that I wasn't happy in the class, and neither of the two professors said anything to me about it, even though it was only 10 [other] people." This might include establishing and reinforcing positive norms for community work and engaging in frequent check-ins with student groups.

4. Where possible, advocate for smaller classes. As Leia pointed out, smaller classes and recitations give instructors and peers the opportunity to get to know each other better. Smaller classes may make it easier to attend to students' self-efficacy, pay attention to social dynamics in the classroom and ensure everyone has positive experiences. Similarly, consider establishing or supporting study groups and affinity groups in your department. Paulette proposes: "Women study groups? Bring your homework and someone who's taken the class before will help you. I think for physics that would definitely be helpful. Or at least, like, a networking kind of thing? It'd be nice to know that there's a woman that has done it before."

These interviews shed light on the experiences of two undergraduate female students majoring in chemistry and physics to give a glimpse of how societal stereotypes can activate stereotype threats [19-21] and how, without adequate support and even in the absence of discrimination, women in physics courses can develop low levels of self-efficacy and be made to feel isolated. As educators, it is critical to reflect upon the experiences of these women since they can impact students' classroom experiences, grades, persistence, and career choices. Physics teachers can benefit from reflecting on these students' narratives in order to create equitable learning environments in which all students can thrive and excel.

4.0 Views of Female Group Leaders

4.1 Introduction

Introductory physics labs can play an important role in helping students learn physics through experimental techniques and the analysis and interpretation of data as well as supporting the development of their problem solving, reasoning and meta-cognitive skills. Many recent studies have focused on the effectiveness of different types of introductory physics labs in a variety of different settings and contexts [28, 88, 101, 223, 229, 262, 278].

However, past studies such as these have placed less emphasis on the experiences, knowledge, values, expectations, and goals of the students who are enrolled in the labs. Cognitive science tells us that it is important to understand what students know and value if we seek to design effective learning experiences for them [6, 221]. Students are the ones who know what they are experiencing and learning vis a vis the goals of the lab course and their own expectations [24, 166, 183, 246, 265]. They also have firsthand experience with the extent to which the learning environment in the lab succeeds in aligning lab goals with their expectations and supports their growth as a physics and science person, i.e., as a person who can excel in physics and science in the lab context. Standpoint theory in sociology suggests that students who are experiencing challenging circumstances are in the best position to explain the nature and cause of their difficulties and how to improve the circumstances [123]. Thus, as a theoretical framework for this investigation, we combine the cognitive perspective that understanding student knowledge and values is essential to designing learning that aligns course outcomes and the goals and expectations of students with the sociological perspective that university students are best positioned to articulate the knowledge and values to drive this process.

We focus on the question of "effectiveness" from the point of view of students in order to understand the alignment between students' perspectives and course learning goals. We view lab learning experiences as effective if they help students move toward their personal goals and/or toward the course goals associated with the lab. The goals for the course include improving students' understanding of physics concepts, learning skills associated with experimental physics, developing the ability to collaborate and communicate physics ideas, and refining students' critical and scientific thinking skills [1, 169]. Students' goals for the course included getting a good grade, preparing for the MCAT exam, engaging in personally meaningful physics learning activities, and collaborating in an agreeable and productive way [76]. To illustrate the ways in which we used the term "effectiveness" during the interviews, consider the following question from our bank of questions: "What are your thoughts on the effectiveness of these labs? How do they compare with other labs you've done in college or high school?" Here, we are asking students to compare their labs with an eye on how helpful the lab was as a learning experience, as seen through the eyes of a student who experienced the labs. Thus, the question uses the term "effectiveness" to probe the alignment between the student's experiences in the lab with their perception of the course's learning goals.

Guided by our theoretical framework, we conducted an ethnographic study in which two researchers observed introductory physics lab courses and also conducted individual interviews with students at a large research university in the US. In particular, individual hour-long semi-structured interviews with 18 students were conducted to understand their views about labs so that they can be taken into account to make the physics labs more effective. Few prior studies have used in-depth semi-structured individual interviews with students to obtain feedback on their views on the introductory physics labs [147, 275]. The interviews focused on diverse issues including the roles of their male lab partners, other peers and the TA in their learning in the physics lab, their views about learning physics in lecture and lab courses, the role of physics labs in promoting conceptual understanding, learning in physics and other science labs, and the role of bio-inspired labs on their learning. Since the physics lab is a collaborative learning environment, the nature of effective collaboration [52, 206, 212] is a thread that runs through the interviews.

This chapter reports on a sub-set of those observations and interviews. We focus on the experiences of women who demonstrated a highly level of agency and acted as group leaders or project managers (or 'Hermiones', to use terminology from [74]). There are three reasons why it is valuable to focus on the experiences of this group of students. First, we note that these students form a unique group whose voices have rarely been heard in a physics context. The design of physics learning experiences tends to focus on the needs of average or 'typical' students, and thus fails to account for the needs of women who are group leaders. Standpoint theory, as part of our theoretical framework, suggests that we ought to listen to these students, in particular, since their needs are not being adequately accounted for in traditional lab instruction. Second, in our ethnographic study, we noted that gender differences in student group work were quite salient. Women often took on certain roles and carried out group work in physics classes in a manner that is not equitable or inclusive, perhaps due to societal stereotypes and biases about who belongs and can excel in physics [79, 161, 197, 260]. This type of gender dynamic in the context of physics has been referred to as "doing gender" while doing physics [61, 115, 302]. We focus on the women who are group leaders because they "do gender" and physics in a way that allows them to uniquely maintain a high level of agency as learners. This is an example of gendered task division in the physics lab [62, 74, 141, 237]. Finally, we choose to focus on the experiences of women in physics because of the many additional challenges faced by women in physics, which have acted in some cases to decrease women students' self-efficacy [157, 205], identity [155, 156, 201], and learning outcomes [2, 30, 116, 127].

Thus, in line with our theoretical framework, this study seeks to understand the experiences, values, expectations and goals of women who acted as group leaders in an introductory physics lab. Specifically, we seek to address the following research questions:

1. How did these students perceive their (possibly gendered) interactions with their lab partners, other students, and the teaching assistant?
2. What types and forms of learning experiences did these students value in the physics lab?
3. How did the physics lab compare with their other lab courses?
4. What elements of the physics lab were engaging, interesting, or effective for these students?

4.2 Methodology

In order to investigate students' introductory physics lab experiences and interactions, we first conducted ethnographic classroom observations discussed elsewhere [74] and then followed those observations with semi-structured interviews [58] with students in traditionally-taught introductory physics labs at a large research university in the US. The investigation was conducted by the author and his supervisor, each with more than a decade of experience as a physics educator. Throughout the investigation, the two collaborated extensively to plan, conduct and analyze the classroom observations and share reflections. For the portion of the broader study we focus on here, the participants for the semi-structured hour-long individual interviews were four female pre-medical students with high agency who the author observed to be group leaders throughout the semester in a traditionally-taught (highly structured) introductory physics lab course for bio-science majors and who strived to ensure that their group did well in the lab.

Since some introductory physics labs are integrated with lecture courses while others are taught as separate courses, we note that our participants were students enrolled in a stand-alone introductory physics lab. This semester-long introductory lab requires the second half of a two-semester introductory physics course as a co-requisite and is often taken in the third or fourth year of students' college study. The majority of students who enroll are bio-science majors with an interest in health-related professions. Approximately 55% are female and 45% are male students. The labs are run by postgraduate student teaching assistants (TAs), who are also responsible for grading the lab work. enrollment is capped at 24 students per lab session and there are several sections running simultaneously in the same semester. Students are graded for completion of their work and, apart from a post-lab exercise, partners receive the same grade for in-class activities. The introductory physics labs have a reputation for being easier than other labs typically taken by students in this course such as the labs for introductory biology, introductory general chemistry or organic chemistry. Students who attend all 12 introductory physics lab sessions typically receive at least a 'B' grade and most receive an 'A' grade. Students in the physics lab worked in groups of two (or sometimes three) to complete the lab procedure during a 3-hour period. Our observations suggest that students self-select their partners at random on their first lab session, as they sit down at an open lab bench. The exception is a very small number of students who partner-up before they arrive in the lab. However, we generally see no significant differences in how these lab partnerships operate. Once formed, groups generally tend to stay together unless the TA requires a re-shuffling (not relevant for the investigation discussed here since there was no re-shuffling for students involved in this study).

Based on the ethnographic study described elsewhere [74], we identified students whose perspectives and experiences would include a cross-section of students who enroll in these types of lab classes, and who could articulate their experiences in individual interviews. These interview findings then would be valuable in order to improve the lab. Approximately half of the students who were invited to participate for hour-long individual interviews volunteered to be interviewed. Since our ethnographic observations suggested differential gender effects in group work in the lab with negative impacts on women, we interviewed more women than men. In addition, we invited more students from mixed-gender partnerships since there was more inequity observed in our ethnographic lab class observations in these groups than in same-gender groups. In particular, we observed that students who worked in same-gender groups tended to collaborate with each other more effectively and equitably. Therefore, in our broader study, the pool of 18 interview participants included 13 who identified as female and 5 as male students. This research was carried out in accordance with the principles outlined in University of Pittsburgh Institutional Review Board (IRB) ethical policy with approval number/ID IRB: PRO15070212. All participants provided consent for this research and the names are pseudonyms to protect their identity.

Here we zoom in on the voices of four female students who acted as group leaders in the physics lab. All of the four pre-medical female student leaders who participated in the individual interviews discussed partnered with others in the physics lab early in the semester and stayed with them throughout the semester. All four were US citizens. In terms of race, one female student, Zara, was Black (she had two lab partners, one of whom was a male and another was a female student), two female students, Kamala and Janet, were Asian (Kamala had a White and an Asian male as her partners and Janet had a single White male as her partner) and one female student, Natalie, was White (with a Black male student as her partner). All students' names are pseudonyms. We note that for these pre-medical students who are intending to go to medical school, physics is a required course and materials related to physics are part of the medical entrance exam (physics makes up 25% of the exam) that students take that is weighed heavily in medical school admission, alongside other requirements [133].

Drawing on our observations, we assembled and refined a list of potential interview questions via an iterative process between the two researchers to serve as our interview protocol. These included questions about the student's background and prior lab experiences; interactions with their lab partners, other students and the TA; thoughts on the structure, mechanisms, and effectiveness of the course; and experiences with task division, including gendered division of labor as well as their views of learning physics in lecture vs. lab courses, the role of physics labs in promoting conceptual understanding, learning in physics vs. other science labs and the impact on their learning of labs that included biological and medical applications (such as blood pressure, a model of the eye, and an EKG). Despite the long list of questions, we made these semi-structured hour-long individual interviews conversational in nature to give interviewees the opportunity to express themselves freely, dig deeply on critical issues, and remain comfortable. Most participants required little prompting and were very keen to share their thoughts with us openly. All interviews were audio-recorded and transcribed.

4.3 Results

Before delving into the feedback received on various issues, we emphasize that the four interview participants whose views we elaborate upon below had all assumed leadership roles and were taking full responsibility for ensuring that the task got done in the lab each week. In particular, they took initiative to bring back to task those in their groups who were distracted by non-lab related things and delegated appropriate tasks to them throughout the lab to keep them engaged. However, our ethnographic observations suggest that because these women leaders spent a significant portion of their time managing their groups' lab projects each week and making sure their lab partners stayed on task, they themselves were often prevented from having the opportunity to actually dive into the lab activities in-depth; e.g., they did not get adequate opportunities to tinker with the apparatus in the lab.

4.3.1 Views About Their Physics Lab Partners

While these female student leaders recognized that if they did not take on the leadership role, things would not get done, overall they found the lab experience to be more positive than their experiences in their large introductory physics courses in which they felt invisible and had even greater self-efficacy issues. Thus, their stressful leadership role in their group in the physics lab at least made them feel somewhat more valued than their experiences in their large introductory lecture-based courses.

Elaborating on the interactions with the two male students in her group, Kamala noted that "the lab is set up in such a way that you're forced to interact with other people, which is good" and added, "but in my group it's a weird dynamic, where the two of them will be unfocused and playing solitaire or something and I'll try to be doing the lab." Then, describing her interactions with her lab partners further, she stated, "In my group it's my responsibility to know what's happening next usually... They just expect me to know what's happening but that's fine..." She also noted that both her male partners claimed that she was better at physics than them but she does not think so because "it just comes down to, are you reading? You should be reading it [lab manual] before.." Thus, Kamala discarded the idea that she was better at physics than her lab partners and attributed her ability to get the lab work done successfully to her efforts and preparation before coming to the lab. Similar themes were echoed by the other women leaders.

Also, interviews suggest that none of these women felt that they had developed a good grasp of physics either from their lectures or from the lab. In fact, despite their leadership role in their group, interviews suggest that these women often had lower physics self-efficacy than their male peers and let their male partners do most of the tinkering with the lab equipment. For example, explaining why one of her lab partners does the majority of the tinkering Kamala noted that her impression was that he was extremely adept at handling equipment and, therefore, he did well tinkering even without reading the lab manual. Similarly, expressing a gendered view about physics learning, Zara said, “When you’re first getting into it, you do look around at your other guy friends or whatever, and there’s maybe a subject that they understand and you’re just not getting it, and you feel like maybe this isn’t something for me.” Zara also admitted that in physics 1, which was a lecture-based course, she felt even less confident than in the lab, stating “it felt like a lot of the other people in my class were getting it and understanding it. And I wasn’t.” She felt better about her standing in the physics lab due to her leadership role than in the physics lectures, acknowledging, “I’m kind of the leader of the group. If I wasn’t there, it would be a detriment to the rest of the group.”

Zara also confessed that she had taken on the role of the group leader because of lack of interest from her physics lab partners stating that they were often not prepared for the lab and it became her responsibility to understand the lab and make sure her group completed it. Then she admitted that “...I think those aspects made it difficult” for her. However, she then added, “I feel like in some ways I enjoy the lab more than my partners. So to get through the lab and to understand it and do it, I end up delegating I guess...” She further elaborated on how she motivates her often unfocused lab partners to do the tasks that she delegates to them by reminding them that if they finish the lab early, they can get out before the end of the three hour long lab period, stating, “I think both of them want to get out of the lab...so if I don’t say, come on guys, let’s go, it will just prolong the whole process of doing it...”

Zara also had some suggestions for how instructors may be able to increase individual accountability in physics labs, saying, “I was thinking, maybe, I don’t know how well this would work, if it was a group of however many, if each partner had to do their own part. You know what I’m saying? I don’t know if this would end up happening, but if each partner had to be the leader for each part, that would get more people engaged. But I don’t know how I would implement that into the groups without one person being like, oh, I understand it I’ll do all the leadership...”. She also wondered whether physics labs could potentially make each student more individually accountable by reflecting on how each person was responsible for certain things in her anatomy lab: “In anatomy, we learn about all the different muscles.

In the lab practical, we have models and we have to identify muscles, where they start and where they end on the body. So if there was a lab practical for a physics lab, you would take a section of the physics lab, [e.g.,] circuit thing, and each student would have to show you how to do one of the circuits, direct or..."

Natalie's views about her lab partner were similar. Reminiscing about her interactions with him and why she ended up becoming the group leader, she said, "The way I understood him, he was just there to get it done and leave..." She then added that because he did not put in the effort, she had to step up her contributions in each lab in order to make sure the lab was completed appropriately. Natalie further noted that her lab partner was totally against spending any time on optional lab activities that would have helped them with better conceptual understanding of the underlying physics. She stated that, since she wanted to develop a deep understanding of physics, at the beginning of the semester she did optional analysis questions in the lab manual but her lab partner was upset and said that there was no reason to do those and so she reluctantly gave up on that part. Elaborating further on her partner's attitude, Natalie said that it was not just her partner who did not care about learning in the physics lab and emphasized that this type of attitude was typical of students in the physics lab particularly because there was no individual accountability. She stated, "I know there were a lot of people that had that same idea. They just wanted to get in, get it done with, and they weren't really getting much out of it."

Janet's reflection about her lab partner also falls in the same category as the other interviewed women leaders. She said, "...he isn't too involved, and I have to push him to do it." Janet also stated that she cared about learning physics in the lab. She also tried to justify her partner's attitude towards the physics lab by confessing that many students in her class don't care about the physics lab particularly because of the way it is structured and incentivized, saying, "The attitude of students in the class about physics lab, it's more so about just finishing the lab and less so about actually learning something." She further elaborated on the students' general attitude in her physics lab, saying "Oh, let's just get these numbers and get out of here. So no one's really understanding, if this circuit is connected improperly, how can we fix it? People just want answers. They don't want to think about it themselves."

Janet further reminisced about her relationship with her lab partner who wanted to do the minimal amount of work and who insisted that they ask the TA about every question that came up as soon as the opportunity arose so that they did not have to think about anything. She said that he insisted that this way they could get done more quickly with the physics lab without needing to apply themselves, stating, "My lab partner and I have a different dynamic. But then I would push him to be like, okay let's try this. It would be better for us to just figure it out on our own rather than just ask for answers." Janet also felt that the incentives for individual accountability and learning physics were very low in the physics lab, which dis-incentivized students from applying themselves. She also suggested that group work in the lab would work well if "they also want to take something away from this lab, to learn something so they can apply it somewhere else. It's hard to get that mentality in other students. You know what I mean?" Individual accountability is indeed important if we want all students to engage in positive inter-dependence and make an effort to learn from the group work.

Janet also noted that her physics lab partner at least did things when she asked him to do something and that, in comparison, she had much worse experiences in other past labs, stating, "physics lab is not the best, but it's also not the worst. I've had worse experiences [in different lab courses] because my partner wouldn't do anything". Then she reflected on her views of what a good lab partner would be like saying, "I feel like a good lab partner should be someone who has the same goals as you with regards to what they want to get out of the class."

4.3.2 Views About Peers in Other Groups and TA in Learning in Physics Lab

Interviews suggest that these female student leaders in the physics lab felt that working with other students not just in their group but in other groups often worked even better than asking their TA for help in clarifying their doubts and it greatly reduced their difficulties. They felt that not only were their peers in other groups more readily accessible but since other students had just encountered those difficulties recently, they often understood their difficulties better than the TA. Also, other students' explanations were often better than the

TA's since they explained to them at the level they could understand them. For example, Zara noted, "Just because the TA is not always around, especially if the lab is a little bit more difficult, it's hard for him to get around to all the groups when we all have questions. So working with the other group is really helpful...I think hearing it from a student, they're on the same level as you, they understand where you're coming from, so it definitely helps to hear an explanation from a student instead of a TA." Then she added that she was always happy to reciprocate, stating, "If they're struggling with a part that we've already done then we'll help them, and it's back and forth."

Zara also stated that she often liked to double check with peers in other groups and get their perspective on the experiment they were conducting each week, stating, "It's definitely helpful to get another perspective, especially if my partners also don't understand what we're doing or the part of the lab that we're doing, then the other group always happens to know something that we don't. And sometimes we'll know something that they don't... I think it's really collaborative, and even better in that we're both trying to figure things out." Encouraging students to take advantage of the expertise of peers in other groups could further motivate students to engage in such productive discussions and help students from both groups [52, 206, 210, 212, 267].

Regarding their TAs, these female student leaders generally had positive things to say; however, sometimes they felt that the TAs were not adequately prepared. In particular, they sometimes found the TA's instructions to be confusing and preferred to ask their peers for help. For example, Natalie felt that the TA did not have a very good understanding of the lab and he did not even think about and try the lab out carefully before the lab like he should have. Expressing her concerns, she stated, "He seems to have an understanding of the lab generally but not in depth." Then she added how she would prepare better for teaching the lab if she were the TA, stating, "If I was teaching the physics lab then maybe I would do the experiment before and go through all the steps. That way I would understand what my students would be doing, and if they got stuck on something then I would have done it beforehand, so I know what should be right..." Making sure that TAs have actually thought about and tried the experiments ahead of time before the actual lab class is indeed important.

4.3.3 Lab and Lecture Courses

These female student leaders who were interviewed explained how the lab and lectures were different. They pointed out that just because someone was good at concepts learned in the lectures, it did not mean they would also do well in the lab and vice versa. They were also generally disappointed that the physics lab course did not help them develop a deeper understanding of the physics concepts, something that was also very important for their medical entrance exam they would be taking in the future.

For example, expressing such views Kamala noted, "Learning the theory is very different from showing up and seeing the equipment. I learned about what a spectrometer does from general chemistry. But I never learned what a spectrometer is and how to set it up. Those are two different things, right? Like, we've seen pictures..." Then trying her best to connect what she did in general chemistry to the pertinent underlying theoretical knowledge, she said, "but my point is... the machine I was talking about gives mass-to-radius ratio of the peaks, right?" Interviews suggest that there may be some confusion pertaining to the difference between a mass spectrometer that separates particles with different masses and the kind of spectrometer students were using in the physics lab to split white light into a colorful spectrum with different wavelengths separating out. It may be useful for physics instructors to explicitly discuss with students that different types of spectrometers may produce spectra of different types of things.

Giving another example from her physics lab about how theory and experiment are very different, Kamala said, "I've known what a parallel circuit is since 11th grade. But setting it up, sometimes, having that conceptual understanding, it looks so different than it does on paper. When you set it up there are so many wires. I was like, I don't know what's going on." Then expressing her frustration at the fact that the physics lab did not even help her develop confidence in being able to set up a parallel circuit, she added, "But I think at the end of it I should be confident I can set up a parallel circuit and I don't think I am regardless of how much theoretical understanding [I have] and how many problems I can solve."

Kamala expressed disappointment about the fact that the physics labs were such that they did not provide incentive for or focus on conceptual understanding of physics related to the experiments and, instead, were mainly focused on carrying out apparently-meaningless procedures. For example, she said, "like in optics, you can figure [it] is a diverging lens, a converging lens, where's the focal point? Then when you do this experiment you should understand why an image is flipped versus why it's not flipped, right? But if that connection's not strong enough, the lab is pointless, because it feels like you're just following a bunch of instructions, and you're writing a bunch of stuff down...". Kamala also felt that the focus on making connections between experiments and concepts that was lacking in her physics lab is more important for physics than for the other sciences that she is familiar with, stating, "With physics especially, that's so important because so many of the concepts are very abstract... In chemistry it's easier because you do chemistry for longer, especially if you're premed, we take chemistry in high school. So we're more comfortable with the subject in general. And because it's less particle-based, and you can see a color change or you can feel heat, so you can feel things going differently, whereas in physics...even in my lab partners I can see that. They might do okay in exams and stuff, but translating that is harder in physics than in chemistry for most people." These types of concerns from students about carrying out meaningless procedures and not getting much out of the physics lab should be taken seriously.

Out of all the four interviewed female students, Zara appeared to be most guarded about not blaming others for her not learning satisfactorily in the physics lab or lecture courses. She described how labs were different from lectures and the inevitable frustrations pertaining to getting things to work in the physics lab, noting that, "I don't know how to prevent the frustration of doing the labs, because it's something you have to figure out. But you can give guidance to an extent, you can ask questions and stuff, but at the end of the day it's trial and error. You have to do it and see if it works or not." She also explained why she liked the physics lab much more than physics lectures, stating, "I definitely think the lab is my favorite part because, like I said, it's a lot more collaborative, it's a lot more hands-on. The stuff you learn in lecture, or even just in the lab manual, it comes to life. It's like, yeah, physics can be fun... I think it's interesting because compared with before where I didn't know anything about physics, now I'm helping my group navigate it." It is clear that despite the burden of being the project manager for her group, Zara takes pride in being the group leader and has a higher self-efficacy in physics lab than lecture.

Natalie also felt that she liked the physics lab better than physics lectures, stating, "I know that I benefited from that because I don't do well in lectures. I tend to space out sometimes... But with the hands-on portion I was able to interact with it so I got a better understanding with it." However, she also expressed frustration at the fact that there was no incentive in the lab for reflecting on and making connections with physics concepts, stating, "Most of the time it would be, okay, this happened, just write it down. And I didn't have a lot of time to really integrate with that information." She also felt that it would be good for the TA to be involved in helping students make the connection between the experiment they were doing and what it meant conceptually: "Maybe if the TA somehow integrates it, explains it to us in more of a conceptual base, that might help us understand the lab..." Natalie then reflected on her introductory physics lecture class and stated, "I understood a lot of concepts, I didn't really learn as much as I could have..." She also felt compelled to contrast the traditional lecture approach with how she learns quantum on her own, a subject that she is very interested in, stating, "The way I learned quantum is, I hike a lot. So I was hiking with a bunch of my friends. We would just sit and discuss these concepts and string theory and talk about quarks and leptons and stuff like that. We would just talk about it because we were interested. And we would throw different ideas and throw different hypotheses: big bang theory, why the universe expands, wormholes. And it was very conversation-based. We didn't do a lot of calculation. It wasn't talked-at you. It was a conversation. It was really cool." Reflecting back on her introductory physics lecture courses again, she recapitulated why she found them ineffective, stating, "But with the lecture it was almost dry sometimes. It was, you have a ball falling out of a plane. It's going to do this because of gravity...I understand the concepts very well, I didn't understand the equations very well and why they integrated in that." Thus, Natalie emphasized that she would have liked more meaning-making opportunities both in physics lab and lecture.

4.3.4 Understanding Physics Concepts

As discussed in the previous section, these female student leaders who were interviewed, valued learning physics concepts and felt that the physics lab fell short of helping them develop a deep understanding of the underlying physics concepts. They felt that the lab was too focused on plug-and-chug approaches and getting the experiments to produce the data and then plugging the numbers into the physics equations to get an answer. For example, Kamala noted, "The lecture gives you the actual equations and the calculation background. But if we come and see this random machine... If the point of this lab is to get us to problem-solve and troubleshoot, then I think it's definitely doing that. But if point is to get us to understand how the physics is applying to the machine, I don't think that's necessarily happening." Kamala was frustrated by how they were spending all their physics lab time getting the experiments to work but there wasn't any incentive or focus on understanding what was actually happening and why. Venting her frustration with the disconnect between conceptual understanding and what was happening in the lab, she described her last physics lab experiment: "In the last lab, I turned the plate 90 degrees and the slit is here, and in the lab it says, now the rays are parallel to what you're seeing. But why is it parallel and how is this set-up making it parallel? I probably couldn't tell you. Does that make sense? Because there's a disconnect between me knowing what a diffraction grating is and me knowing what the speed of light is and me seeing this machine and knowing how turning something changes."

Natalie returned to contrasting how she effectively learns quantum concepts via discussion with her learning in the physics lab by elaborating, "my boyfriend and I just talk about it [quantum] for hours and hours... I like understanding things rather than just observing and writing things down [as in the lab]... I like having the concepts behind when it comes to stuff." She also emphasized how learning concepts makes her feel she really has a good grasp of the physics topic and it makes her feel confident, explaining, "I love understanding concepts, I want to get the most that I can out of it... Once I understand the concepts, I get confident." She then made a transition from the physics lab to her college courses in general stating she always likes to think deeply about concepts in her other courses also but the STEM courses often have too much material thrown at students and don't necessarily give them an opportunity to think and understand. Reflecting on these issues she stated, "I like to apply that to every class that I go to. I want to get the most out of it... And in STEM there's a lot of information thrown at you. Get it done, move on to the next class. Learn it, get it done with, move on to the next class." Then assuring herself that she is not alone and students at all colleges are facing similar situation she said, "Not just at Pitt, at every college I've had my friends go to, I've talked to. It's just a lot of shoving down information. You don't really have the time to understand the concepts." It is important for physics instructors to take these reflections from students seriously and strive to balance the amount of materials covered with providing sufficient opportunity for students to actually understand what they are learning.

Natalie also opened up and described how people she has encountered in life have had very different opinions about her ability to succeed in science and medicine depending on whether they understood her quest to grasp concepts and be challenged or not. She reminisced about the countless times when people did not believe in her ability to be able to do science all the way from K-12 to college and those reassuring times when they actually believed in her. In particular, she said that she asks a lot of questions because she really likes to learn concepts and be challenged. However, many people misinterpret her deep quest for learning and think that she is asking questions because she is not capable of understanding quickly. With sadness in her expression she stated, "A lot of people told me no. I had tutors tell me I could never become a doctor, I could never go into something like neuro. I could never go into... because I wasn't smart enough, when I really just wasn't challenged enough. There's no reason for me to try if you're just going to tell me the same thing over and over."

Then she reminisced about the science classes that really engaged and challenged her and validated her as an individual who has enthusiasm for deep thinking. She felt recognized in those classes and those courses helped her learn to think like a scientist. Recalling two of those classes she said, "There was this one class I took my freshman year, Brain and Behavior. It was my first neuro class at Pitt. And I talked to the professor, she went over something very slowly, she went into depth, she went into all the details. I was able to sit down. I spent the rest of the day coming up with ideas. I went up to her and I started talking with her about all these ideas I came up with, research ideas, with it. And I was really interested in doing that kind of research, there isn't a lot of papers out there about it. And there was another class that I took, it was functional neuro-anatomy. We talked about the visual system, and why something didn't work. And the teacher gave us time to really think about it. And all of us in the class were asking questions, trying to understand it, trying to work with the material, to come up with different research proposals." Then contrasting these thought-provoking science courses with other classes, including the physics classes which she felt were not as thought-provoking, she stated, "And then suddenly we're getting pounded with material left and right...And it just comes in and goes out. You don't really have time to absorb it, think about it, and generate. I think that's a problem with colleges all across."

Janet also made it abundantly clear that she cares about learning physics concepts, stating "Especially things like flow viscosity and things like that, where we look at that with our blood flow. It's so important to our biology. I feel like all sciences merge together."

4.3.5 Comparison of their Chemistry and Biology Labs with Physics Lab

All of these female group leaders in the physics lab felt that their chemistry labs were the most effective in helping them learn concepts and helped them integrate whatever they learned in the lab with the lectures. They also liked the fact that their chemistry and even biology lab's main focus was not on manipulating lab equipment, collecting data and plugging them into equations to spit out answers, which seemed to be the main focus in their physics lab. A mismatch between their expectations and the actual focus of the physics lab often seemed to frustrate them about why there was such a lack of focus on conceptual understanding in their physics lab.

For example, describing her chemistry lab and comparing it with the physics lab, Kamala noted, "A lot of chemistry is knowing what reagent to add in excess, and what doesn't need to be added in excess. And they'll explicitly state, i.e., the lab TA, before we start the lab, why are we adding this in excess? Because XYZ reason. Or why are we... even there, maybe there were some points that were lost in translation but even there I felt like I understood more of why I was doing something than I did in physics lab."

Reminiscing about a physics lab experiment that she already had a good grasp of before doing it, Kamala stated, "...there's some labs that are more obvious to us. Like the eye one, there's a lens. Probably because I've done a similar thing in high school, so this was the second time around, so it's more likely I'd understand. But it was more apparent to me what was happening. We were pumping fluid into the eye, so we could see the lens changing shape, right? So if you know enough conceptual things, like this is a lens, you're more likely to understand why you're doing something." She contrasted it with physics labs in which she had great difficulty understanding anything due to her unfamiliarity with relevant concepts before doing the lab and the fact that no incentive or support was provided to understand them. For example, about the diffraction grating lab she recently encountered, she stated, "Especially when we encounter a machine like we did this past week, there's really nothing in your past education that shows you what this machine is supposed to be doing. At least for me, it wasn't necessarily clear what we were supposed to be doing. There was a thing you were looking through, and there was another tube, and an angle you were changing. I get that light is hitting this plane that is reflecting back at us, but I think that clarity about knowing this is what's happening was not there. It was kind of magical, right? Like, we turned it and... somehow we got a diffraction grating. But I couldn't tell you why exactly I was doing each step."

Kamala also had similar views about her chemistry and physics lab reports and stated, "Our lab reports in chemistry are more conceptual. The questions we're asked are more like, why do you do this? ... It's not like, take this data and do a bunch of calculations and give us numbers. That's what our lab reports are like in physics." On the other hand, about her physics lab she felt, "because the evaluation is not conceptual, we're less likely to push ourselves to understand why something is happening".

Zara recalled that her bio labs motivated all students to understand concepts well by requiring poster presentations, stating, "The other labs that I've been in are SEA-PHAGES in the biology department... At the end of those labs, the end result is producing a poster. In that sense, everyone has to understand what's going on because you're presenting it to other people. When you have the poster presentation, one person does one section, and another person can do another section, but they also encourage students to understand the entire poster themselves, too." Physics departments should also consider giving similar opportunities to students to present in poster format to their fellow students and instructors.

Similar to Kamala, Zara also emphasized how chemistry labs focused on concepts, which she found to be very useful. Reminiscing over the pre-labs assignments in different labs and how the physics lab was different she stated, "In one of my gen chem labs... you would have to explain things more. Whereas in physics it's like the pre-labs are, like, using equations to figure out an answer... In the gen chem labs, it's going into detail about how you explain the stereochemistry or how you explain... it's like applying concepts. So even if you didn't understand a certain mechanism, they'd take that concept and put it into something else for you to explain, if that makes sense?"

Natalie talked about how much she enjoyed and benefited from her research-based labs in chemistry: "I really like the research-oriented ones because it pushed me to really use the material I learned in a novel way. I'm applying to something that not a lot of people have seen what happens. You can't predict the results. You really have to have a strong understanding for what's going on in order to apply it." She particularly appreciated the fact that these chemistry labs were asking students to apply concepts to situations they had never encountered before in any courses, stating, "...it's really trying to push past the boundaries of what you're taught in the classroom, because you have to apply things and come up with new things in order to understand what's going on."

Recalling her chemistry lab Janet said, "those labs helped us figure out, oh, that's how you do the calculations based on those numbers based off this data we just collected." She lamented that in physics lab she does not feel the same connection between the experimental work she is doing and the concepts. It seemed to her that they are just collecting data with one equipment after another and churning answers using physics equations without understanding the concepts. However, Janet acknowledged that she appreciated the few times she was able to discern how physics and chemistry labs connected with each other. For example, she found it revealing that those labs used spectra and spectroscopy in different ways, stating that in the physics lab "seeing the light and the different spectrum. It's very different from the spectroscopy that we see in chemistry. So I thought that was pretty cool. Seeing a similar machine but in a different way." Thus, similar to Kamala, Janet was excited about having spectroscopy used in both chemistry and physics contexts. Physics instructors can take advantage of these opportunities to help students reflect upon the underlying similarity and differences between different types of spectroscopic devices/measurements and have students discuss the connections between the chemistry and physics contexts students have encountered.

Janet also expressed great enthusiasm for the connection between her bio lab and research she was doing with a professor, "I was working with RNA primers... DNA replication, qPCR. These are things I learned in class and I was like, oh I'm seeing this in real life. And that's what the numbers look like, and that's what they mean in regards to what I'm learning in class. And I thought that was really interesting, to see that connect is really great." She was really excited about being able to interpret data and make the connections, stating, "Oh, this is what's happening. I was able to see. I got numbers from qPCR, for example, and that tells me this protein is expressed this much in that area of the brain. And so that tells me those cells in that brain area are being regular to produce this much protein. And then, from that I was able to say, oh, we can see what's being down-regulated/up-regulated in different people. And those concepts kind of connected back to what I learned in foundations of biology. Oh, I learned about DNA replication, I learned about how we go from DNA to RNA to protein. And I learned about how the amount of RNA you make affects the amount of proteins you're going to make. I could see all those concepts connecting to what I was doing in research. And that was very meaningful to me because I was doing something I learned about." She was disappointed that physics lab did not provide such opportunities.

4.3.6 Interest in Bio-Inspired Physics Labs

All of these female student leaders noted that they appreciated being able to see the connections between biology and physics via biology-inspired physics labs. For example, Kamala noted, "I really liked the eye one... it's a direct application between optics and light. You can see the lens getting big. You are literally pumping water into it with a syringe. I thought that was very cool but I'm biased because I'm premed... We did an EKG thing and ECG. That was cool too... I'm in human physiology... I was really excited. PQRS waves! I know what valve was closing too! I knew the physiology behind it, so seeing an EKG was really exciting! I think I did it on my partner, or we were doing it on our partner, and it was fun."

Kamala also said that she would have loved to do more bio-inspired physics experiments that were not among those she actually got to do, "like learning the mechanics. This is still torque, me moving my arm. But doing an experiment on that would be cool because that's the foundation... especially when most of the students in this class are pre-health in some way." It would indeed be valuable for physics instructors to take advantage of student interest, e.g., in biology, in developing physics labs.

Discussing how much she liked certain labs that focused on bio themes, Zara said, "Learning about the eye kind of overlaps with the anatomy that I'm interested in. Seeing different converging lenses and diverging lenses, how that's involved in physics 2 I think was really cool. And the EKG... it was really cool to see how the heartbeat was looking on the thing. And being able to understand it was really cool too."

Natalie felt that application-based labs would definitely be the best for students who were not physics majors, stating, "I know that a lot of students in that lab are pre-medicine. So by making it more application-based to the world, not just medicine, maybe it might have students get more out of it". Then she described the value of taking a neuro-physiology course at the same time as the physics lab in which she was learning about circuits and how she was able to make the connection between the two course materials: "The circuit labs I found very useful. I'm also taking neuro-physiology right now, which is heavily physics-based, so that helps me apply it to other classes."

Natalie also reflected on how she was trying to make the connections between her physics lab experiments and what she had learned in other courses but sometimes those connections were not easy to make, stating, "Other labs like... the heart monitor was pretty cool, just because it was applicable to me. The EKG... and I know there's a lot of parts because there's a lot to cover around that topic. But something like the heart monitor, I didn't have time to understand why the T-wave was inverted if I put the black electrode on this arm versus the other arm." It was clear from the discussions that providing scaffolding support to students in order to make these types of connections between biology and physics would increase student engagement since many of them really want to understand those connections.

4.3.7 Mechanics vs. Other Physics Lab Experiments

Three of the four female lab group leaders expressed more positive views about the physics lab experiments that focused on mechanics but Natalie was more enthusiastic about the other physics experiments that focused on electricity, magnetism and optics. Kamala noted that without much support for understanding what was going on in the physics lab experiments, she liked the mechanics experiments the most because they were easier than other topics to grasp on her own, stating, "A diffraction grating, you can't see the light, you can't see the rays, you don't know what's causing this. So if you don't know why something's happening, there's really no basis to understand what you're doing. Whereas in mechanics there is. You can see something is falling. You can see something is slowing down."

Janet also noted liking mechanics labs, stating, "I really enjoyed the first lab, the roller coaster... you drop a metal ball. Those are really fun because you see the effect of gravity on different weights of the ball. I think something that could be better is, instead of just testing two balls, what if we tested different materials? Like, how does the material affect the weight and all that stuff, and what could you predict as the distances and stuff like that? That's something that's interesting."

The only student among them who was more enthusiastic about other experiments and did not care much about mechanics was Natalie, who said, "I know the ball dropped, I've observed gravity my entire life. I don't get much out of that lab." In other words, Natalie was more interested in experiments that were novel and different from what she had experienced in her everyday life. Focusing the introductory labs on a wide variety of physics topics would serve students with different interests.

4.4 Summary, Discussion and Implications

Feedback from students who have taken the lab can play a critical role in revamping and designing effective labs to improve student learning. This feedback can help physics departments refine the goals of their lab courses and make them more consistent with what would be most effective for students with a certain background and future professional aspirations. The feedback can also help instructors understand how to frame their instruction and achieve buy-in from learners in a way that usefully aligns course learning goals with the expectations and values of students.

We have used a theoretical framework that combines the principle from cognitive science that it is important to understand students' knowledge and values with the principle from sociology that students are best positioned to describe the challenges they face and potential resolutions for those challenges. Our investigation involved hour-long semi-structured individual interviews with four female lab group leaders who were enrolled in a traditionally-taught introductory physics lab for bio-science majors and who took on the role of project managers in their lab groups throughout the semester. The ethnographic lab observations that the author and his supervisor conducted suggested that these female students who were project managers in their lab groups had full responsibility for managing their groups' lab work and making sure that the lab work was done appropriately each week [74].

We find that there were inequities in group interactions, e.g., some group members such as the interviewed female students stepped forward to fill the vacuum in their lab groups. They became group leaders and had a disproportionate level of project management responsibilities. Although some of them mentioned taking pride in being the group leader, explicit efforts should be made to make the group work more equitable in physics labs so that one person does not have disproportionate amount of burden for making sure everything was done each week by the lab group [74, 141, 237]. One suggestion to make the physics lab equitable that came from the interviewed women was increasing individual accountability and making sure that at least part of the lab grades were assigned based on each individual's effort and understanding instead of collectively for each group. Also, regularly changing group composition may help to reduce certain types of inequities in the group work that was observed. An often-recommended approach is to assign and rotate roles within each student group. Another important recommendation is to not isolate minority students, e.g., one woman in a group with two or three men, since these types of situations make women particularly vulnerable to taking on an inequitable gendered role [126]. Providing both individual accountability and support for collaborative strategies that empower all students to contribute equally could improve the effectiveness of student collaborative work [52, 267] in the physics lab.

Furthermore, due to the aforementioned societal stereotypes about who belongs in physics and can excel in physics, fixing the gendered nature of the physics labs we found in our investigation may also require long term efforts [220]. For example, increasing the representation of women in physics particularly in leadership positions, e.g., in the form of female TAs and professors, could be helpful [23]. When asked what can be done to improve the physics self-efficacy of women like her, Zara said, "maybe having more women in science, more representation with women professors. Like, seeing yourself in the professors, I think that would definitely help with other female perspectives going into physics class." The self-efficacy of women in physics lags that of men, and is an important predictor of academic retention and success [20, 124]. Zara's response is revealing and shows how historical societal stereotypes and the lack of female role models in physics impact women like Zara even in physics courses in which they are not underrepresented (e.g., her physics lab for bio-science majors had 60% women).

Although this study focused on a lab that was 55% women, we note that our ethnographic observations of labs that did not include pre-medical students (and thus enrolled less than 50% women) showed similar patterns. In such labs sections, we saw women take on the group leadership roles [74] and have similar experiences to those described by Janet, Kamala, Natalie, and Zara. Our observations suggest that the findings we describe here primarily arise from the culture of physics, and not from the specific context (i.e., an approximately-even gender ratio) in which these four students studied.

The interviews also suggest that the supervising TAs were not always well prepared for the labs or had not adequately thought about and tried the experiments that their students would do each week. Therefore, providing good professional development of the

TAs who run the labs and making sure they are prepared for their teaching duties are critical. What is equally important is getting 'buy-in' from the TAs about how to run the lab effectively and equitably for ensuring that the introductory physics lab is indeed functioning as envisioned [316].

Students should also be encouraged to talk to other students in different groups since one TA may not be able to help all groups at a given time. More importantly, as interviews with these women suggest, since other students have learned the concepts recently, they can often understand other students' difficulties better than the TA and provide more useful feedback and help. However, creating a lab environment in which all students (and not only some students) feel comfortable asking their peers in other groups for help without feeling judged is really important. This can be especially important for women who face the challenge of "doing gender" while also doing physics [115, 302]. The professional development of TAs can again play a key role in creating such an inclusive environment in the lab.

Interviews suggest that the female students typically felt that there was a disconnect between their physics lab and their desire to learn physics concepts, and that the lab was not designed to help them learn physics concepts. These students also expressed that while they recognized that the physics lectures and labs were different, they did not appreciate that the physics lab did not provide incentive or support for helping them learn physics concepts. They felt that the fact that merely being present in the physics lab was sufficient to get a good grade and the fact that the labs were not designed to help students learn physics concepts made many students who already were skeptical about the physics lab even more disinterested and disengaged. The students explicitly noted that they felt that the labs were too structured and procedural. These interviewed students also pointed out that since many of the students in this lab were pre-medical students who had to take the medical entrance exam focusing heavily on physics concepts, lack of focus on physics concepts greatly reduced their level of interest and engagement in the lab. Past research has shown a tight connection between physics interest and self-efficacy in physics [124, 156]. The students also felt that increasing individual accountability in the lab can go a long way in increasing student engagement as well as learning.

Furthermore, the interviews suggest that the physics labs should also take inspiration from chemistry and biology labs that students appreciated significantly more. They found them more heavily focused on helping them learn relevant concepts in the lab context. It is not surprising that chemistry and biology labs' focus on concepts was something that these students aspiring to end up in health professions, who had to take an entrance exam focusing on these subjects, found useful. These interviewed students also pointed out that some introductory chemistry and biology labs required them to do novel experiments that were closer in spirit to authentic research and helped them learn to think like a scientist. This suggests that these students viewed their physics lab as less effective for helping them to learn the critical and scientific thinking skills [1, 169] that were goals of the course. They also felt that some of the labs that required students to present the lab work in the form of posters were effective in ensuring individual accountability in addition to helping students develop the ability to communicate scientifically. In particular, some of them noted that some labs in other subjects promoted individual accountability because, e.g., they required each group member to present a part of the poster to the class but each student was expected to be able to answer questions about any part of it.

Finally, we note that in an era in which interdisciplinary training is more important than ever, thoughtful design of physics labs for bio-science majors can be particularly important for keeping students actively-engaged and providing them with appropriate training. All of the interviewed students noted that they really appreciated the bio-inspired labs and some explicitly noted that they would have liked more physics labs with similar themes integrated in the course. They also wanted the physics lab to help them make better connections between physics concepts and concepts in other areas of science they were interested in. Giving students an opportunity to make these types of interdisciplinary connections should be an important goal particularly of physics labs that are for students primarily majoring in bio-sciences or other disciplines.

Based upon the feedback received from interviewed students, we are developing a new grading rubric that takes into account individual accountability in addition to promoting positive interdependence between the physics lab group members. Rotating group members a few times per semester as well as assigning and rotating the roles of different group members within each lab group are also things we have begun to implement in our introductory physics lab. We are incorporating opportunities for reflections on these issues and activities focusing on effective strategies for physics labs in our TA professional development workshops. We are also monitoring the implementation of the strategies learned in the professional development program by the TAs in the physics lab. Physics departments should reflect upon the consistent views articulated by the interviewed students in order to improve the effectiveness of their physics labs.

5.0 What Makes a Good Lab Partner?

5.1 Introduction and Framework

The introductory physics lab brings together students to collaborate actively in scientific meaning-making [4, 8, 35, 52, 85, 96, 143, 169, 262, 274, 311]. While evidence-based collaborative and active learning may improve student outcomes on average, the impact of these pedagogical strategies on diverse pools of students may not be uniformly positive [44, 197, 246].

In order to understand how collaboration in a primarily traditional introductory physics lab can impact a student's identity trajectory [44, 106, 124], we investigated the ways in which interactions with lab partners affected students' interest and self-efficacy in an introductory physics course. We approached the problem in three ways: first, by asking students about the characteristics of an ideal lab partner; second, by investigating how students' perceptions of the distribution of the lab work between the partners (equal or unequal participation in all aspects of the lab) is related to self-reported changes in their interest and self-efficacy in physics; and third, by assessing whether gender might play a role in the relationship between lab work distribution and self-efficacy. This quantitative work is designed to complement qualitative analysis reported elsewhere [74].

In order to contextualize our research, we adopted a framework based on identity in communities of practice that has been used before in physics education research [51, 149, 244]. A community of practice consists of three elements: the domain, the community, and the practice [300]. For our context, the domain consists of student learning in the introductory physics lab, the community is the pair (or sometimes triplet) of students who collaborate each week on their lab work, and the practice is a 3-hour guided inquiry lab investigation that the students undertake together [278]. In a community of practice framework, students develop their domain-related identity (i.e., their physics identity, or the "type of person" they see themselves as [106]) within a community of practice through engaged participation [300] in relevant activities. In the case of the introductory physics lab, we focused on how students' interactions with their partner(s) affected the development of their physics identity as they conducted experiments and analyzed their results.

One aspect of the process of identity development is 'mutuality of engagement,' a type of competence in which students develop the ability to engage with other members of their community and to establish relationships around their practice rooted in mutual and community benefit [244, 301]. As students do their lab work together (engagement), how do they interact productively, give and receive help, and coordinate the boundaries of their cooperative work? As a relevant example, consider two different ways that pairs of students could split their lab work equally. In partnership A, one student does all the writing while their partner sets up the apparatus and makes the measurements. In this case, if each task takes 50% of the time and effort then the students might view their mutuality of engagement as fair. However, this mutuality of engagement is not equitable because the students do not have equal opportunities to benefit from the various lab learning experiences, such as operating the apparatus and recording and analyzing the data. By way of contrast, in partnership B two students share the work equally: they each get opportunities to carry out all the aspects of the lab, including manipulating the equipment and recording and analyzing the data. In this case, the mutuality of engagement is more equitable.

In this investigation, we focused on two facets of identity development that were assessed using survey data: physics interest and self-efficacy. Physics interest has been associated with both course and career decisions [124]. Self-efficacy (sometimes also called competency belief) relates to a student's belief in their ability to succeed in a certain situation, task, or domain [20], and has been associated with student performance and persistence [205, 253]. Both physics interest and self-efficacy have been shown to impact students' identity development in physics [111, 156]. In this research, we investigated how the nature of a lab-group's mutuality of engagement might have affected the development of students' interest in physics and their self-efficacy.

5.2 Methodology

Participants in this research study were students enrolled in a one-semester introductory physics lab course at a large state-related research university in the USA. These students included physical science and health science majors, but not engineering students who take a different lab course. For three semesters, data was collected via bubble sheets filled out at the end of the lab period, and for the most recent semester the survey was filled out online. Since students received bonus points for completing the survey, the response rate was more than 90% for all four semesters, including the most recent.

In the lab, students worked in pairs (or triplets, if necessary) to complete a 3-hour course of lab work each week. Student partnerships were self-selected and appeared to be mostly random, as students typically worked with whomever they sat beside at the start of the first lab session of the semester. Pre-lab and homework assignments were completed individually, but a small amount of collaboration was typically required outside of the lab to complete and submit the digital lab report. Lab partnerships were stable through the semester and received the same grade for their lab reports.

What Do Students Want? In the first part of the investigation, students were asked to respond to the prompt, "In the space below, please describe the characteristics of an ideal lab partner." We collected data from 120 students during two semesters for this part of the investigation. Responses were read and a generative coding scheme [228] was developed to categorize the responses. Two researchers independently categorized the first 49 responses in order to assess the reliability of the coding scheme and categorization process, producing a Cohen's $\kappa = 0.75$ which indicates 'excellent' agreement [97].

An Even Split or an Equitable Distribution? As outlined in the introduction, there may be a difference between students splitting their work evenly (e.g., partnership A) and dividing the work so that all students get to participate equally in all aspects of the lab, including both work with the apparatus and recording and analyzing data (e.g., partnership B). In order to investigate whether students benefit from participating equally in all aspects of the lab, for the second part of the investigation we collected three semesters' of data from 163 men and 258 women to questions pertaining to peer effect on interest. We also analyzed responses from 300 men and 492 women pertaining to peer effect on self-efficacy on an end-of-semester survey in the lab, including a fourth semester of data in order to increase the sensitivity of our analysis. These survey questions were validated along with other constructs on a larger motivational survey using think-aloud individual interviews with students to ensure that students interpreted the questions correctly, factor analysis of student responses to ensure that the questions grouped into clusters as expected, as well as analysis of Cronbach's alpha for each factor and Pearson correlation amongst different factors [156, 201]. In order to investigate how equal participation predicts peer effect on physics interest or peer effect on physics self-efficacy in a multiple linear regression model, we used gender and a self-efficacy construct extracted from E-CLASS [72, 308] as controls to improve the explanatory power of the model. We note that the larger pool of data (300 men and 492 women) for how equal participation predicts peer effect on self-efficacy allowed smaller effects to manifest as statistically significant if they were present.

We obtained gender information by connecting student responses with anonymized institutional records of students' genders, which were self-reported as Male or Female at the time of application to our university. While we recognize that gender is not a binary construct, other gender identities were not included in this institutional data and all students included in this study selected one of these options.

The set of questions on the survey related to the effect of peer interactions on the participant's interest and self-efficacy in physics is shown in Table 2. We hypothesized that changes in a student's interest and self-efficacy in physics might depend on the nature of the group's mutuality of engagement, and specifically on whether students participated equally in all aspects of the lab, which we assessed by querying whether the survey taker and their partner "participated equally in each component of the lab which includes: (i) manipulating the equipment and (ii) data analysis" on a 5-point Lickert scale from "Strongly Disagree" to "Strongly Agree".

As noted earlier, we also analyzed responses to the E-CLASS survey [308], which students completed at the start and end of the semester, in order to control for students' self-efficacy in our linear regression models. We extracted a subset of four items from the E-CLASS that are related to the participants' own perceptions of their self-efficacy [72], shown in Table 2.

Table 2: Validated survey items related to the E-CLASS self-efficacy, peer effects on physics interest, and peer effects on self-efficacy constructs.

Self-Efficacy ECLASS items
If I wanted to, I think I could be good at doing research.
When I approach a new piece of lab equipment, I feel confident I can learn how to use it well enough for my purposes.
If I try hard enough I can succeed at doing physics experiments.
Nearly all students are capable of doing a physics experiment if they work at it.
Peer effect on physics interest items
My experiences and interactions with other students in this class
... stimulated my enthusiasm for physics.
... made me enjoy physics more.
... increased my interest in physics.
Peer effect on self-efficacy items
My experiences and interactions with other students in this class
... made me feel more relaxed about learning physics.
... increased my confidence in my ability to do physics.
... increased my confidence that I can succeed in physics.
... increased my confidence in my ability to handle difficult physics problems.

Finally, before carrying out the multiple linear regressions, for the E-CLASS self-efficacy, the peer effect on physics interest, and the peer effect on physics self-efficacy constructs, we followed established practices for these survey items by collapsing 5-point Lickert scale items to a 3-point scale (i.e., “Strongly Disagree” and “Disagree” were combined, as were “Strongly Agree” and “Agree”) [144, 308]. Scores for these three constructs were constructed by assigning +1 for agreement with each item, -1 for disagreement, and 0 for neutral, and then averaging over all the items in each construct. The gender variable was assigned as $0 = male, 1 = female$. All factors other than gender were normalized to have a mean of 0 and a standard deviations of 1 so that regression correlations are reported in terms of standard deviations [55].

5.3 Results

What Do Students Want? The results of the categorization of student responses to the prompt asking for characteristics of an ‘ideal lab partner’ are reported in Table 3. Since a single response could include several characteristics of an ideal lab partner, the total number of occurrences of different characteristics exceeds the number of responses categorized.

The most frequently cited characteristic for the ‘ideal lab partner’, reported by more than half of the students, was a willingness to split the work evenly. Responses coded for “fair split” included those that expressed a desire for a 50/50 split, a partner willing to do half the work, or a partner willing to “do their share.” These responses reflect a mutuality of engagement that may not be equitable and that may not provide the benefit of equal engagement in all aspects of the lab to all students, as illustrated by example partnership A in the introduction. Doing half the work is a relatively simple form of mutuality of engagement in comparison with the advanced levels of cooperation, relationship-building, and mutual benefit that are possible when people collaborate equitably and effectively. In the 120 responses, not one described an ideal lab partner as someone who supported equally sharing each aspect of the work, as in example partnership B.

Personality traits that allow students to easily cooperate and get along well with their

Table 3: Characteristics of an ideal lab partner, according to students.

Characteristic	Occurrences
Fair Split	62
Smart, Knowledgeable	31
Easygoing, Sympathetic	28
Hardworking	27
Communicates	27
Prepared	22
Friendly, Kind	21
Helpful	18
Enthusiastic, Cheerful	18
Efficient, Timely	18
Contributes to my Learning	14
Intellectually curious	8
Serious	3
Total number of responses	120

partners, such as being easygoing, communicative, friendly and enthusiastic, were frequently cited as being desirable. Another class of characteristics frequently cited were those that might help students complete their lab work efficiently, such as having a partner who is knowledgeable, hardworking, helpful, and timely. Notably uncommon among the responses were characteristics that might be associated with improving one's learning in the lab, such as the partner being intellectually curious or the partner contributing to one's learning.

An Even Split or an Equitable Distribution? We hypothesized a multiple linear regression model,

$$\begin{aligned}\hat{y}_1 = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \hat{\beta}_3 x_3 \\ + \hat{\beta}_{12} x_1 x_2 + \hat{\beta}_{13} x_1 x_3 + \hat{\beta}_{23} x_2 x_3\end{aligned} \tag{1}$$

to account for how students' responses to the equal participation prompt (x_1) predict their responses to the peer effect on physics interest construct (y_1), using the E-CLASS self-efficacy (x_2) and gender (x_3) as controls. The model also included all two-way interactions between the three independent variables (x_1, x_2 and x_3). In the model, hats denote predicted (or expected) values determined by an ordinary least squares fit in the regression [55].

The multiple linear regression was carried out using an ordinary least squares linear regression algorithm in R [238]. The results (see Table 4) support the hypothesis of a relationship between participants' likelihood of saying they participated equally in the lab work and their likelihood of reporting that interactions with other students improved their physics interest, with a 0.16 (0.06 to 0.26 at 95% confidence [55]) standard deviation increase to the peer effect on physics interest for each 1 standard deviation increase in equal participation. Gender was a significant predictor of physics interest, with women on average reporting a value for the peer effect on their physics interest that was 0.31 (0.13 to 0.49 at 95% confidence) standard deviations lower than for men. Likewise, the self-efficacy construct from E-CLASS served as a useful control. No interactions were significant (at $p < 0.05$ [55]), thus interactions were removed from the model and are not shown. The model accounted for 11% of the variance in the outcome variable, as calculated using the adjusted R^2 [55].

We also hypothesized an analogous multiple linear regression model which focuses on how students' responses to the equal participation survey question predict their responses to

Table 4: Results from a multiple linear regression model focusing on how equal participation predicts peer effect on physics interest controlling for E-CLASS self-efficacy and gender.

Peer Effect on Physics Interest	β	SE
Equal Participation	0.16***	0.05
E-CLASS Self-Efficacy	0.23***	0.05
Gender	-0.31**	0.09
Adjusted R-squared	0.11	
Sample Size	421	

$***p < 0.001$, $**p < 0.01$, $*p < 0.05$

the peer effect on self efficacy construct using E-CLASS self-efficacy and gender as controls. This model has the same form as the previous one (Equation 1), with the peer effect on self efficacy (y_2) replacing the peer effect on physics interest (y_1) as the outcome variable. As in the previous model, all two-way interactions were included.

The linear regression analysis, reported in Table 5, was performed in the same way as the regression in which the outcome variable was peer effect on physics interest. Non-significant ($p > 0.05$) interactions were removed iteratively and are not shown. This model accounted for 16% of the variation in the outcome variable and included a statistically significant interaction between gender and equal participation, while both gender and equal participation individually were not significant predictors of peer effect on physics self-efficacy. An increase in equal participation predicted no change to the peer effect on self-efficacy for men, with a non-significant 0.02 (-0.03 to 0.07 at 95% confidence) regression coefficient. However, the story for women is quite different. Using values from Table 5, women reported a (0.02+0.17 =) 0.19 (0.10 to 0.28 at 95% confidence) standard deviation increase to the peer effect on physics self-efficacy for each 1 standard deviation increase to equal participation.

These results may also be understood graphically, as depicted in Fig. 2. For men, variation in their response to the equal participation question did not significantly affect the outcome (i.e., in Fig. 2 the slope for men is nearly horizontal). However, for women, there

Table 5: Results from a linear model focusing on how equal participation predicts peer effect on physics self-efficacy controlling for E-CLASS self-efficacy and gender.

Peer Effect on Self-Efficacy	β	(SE)
Equal Participation	0.02	(0.05)
E-CLASS Self-Efficacy	0.34***	(0.03)
Gender	-0.11	(0.07)
Gender × Equal Participation	0.17*	(0.07)
Adjusted R-squared	0.16	
Sample Size	792	

$***p < 0.001$, $**p < 0.01$, $*p < 0.05$

is a significant effect: response to the equal participation question is positively associated with the peer impact on self-efficacy (i.e., the slope for women in Fig. 2 is both positive and significantly different from the slope for men). Thus, we can conclude that men who participated equally were not more likely to report that collaboration improved their self-efficacy, but women who participated equally did report a boost to their self-efficacy because of peer interactions.

5.4 Discussion and Implications

In the first part of the investigation, we saw that students focused on issues of social interaction and work completion when queried about their views of what makes an "ideal lab partner". They wanted a lab partner they could get along with, and who would help them to complete the work. Most importantly, though, they wanted a lab partner who would be willing to do their "fair share": 50% of the work.

While a 50/50 split might sound equitable on the surface, as reported by previous work [61, 74, 141, 236, 237], more so than in same-gender groups it is not uncommon for

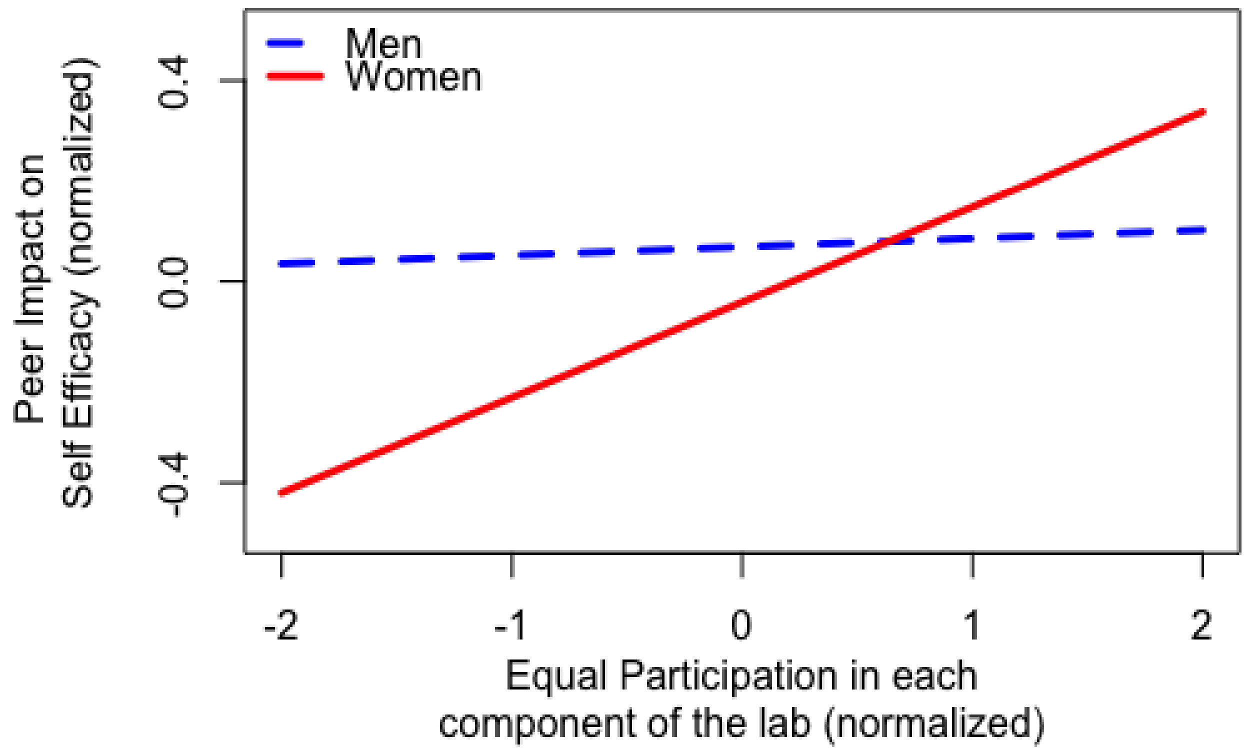

Figure 2: Linear analysis of students' perception of the impact of peer interactions on their self efficacy for men and women.

students in mixed-gender lab groups to engage in gendered task division in which men tend to do one type of work while women tend to do another. Thus, the mutuality of engagement associated with a “fair split” of the work is markedly different from, and may be less equitable than, the mutuality of engagement associated with a distribution of the learning activities in which each group member participates equally in all aspects of the work. It is not just the amount of work that partners must share, but the types of work as well.

The difference between these two forms of mutuality of engagement is shown to be relevant in the second part of the investigation, in which we found evidence that higher mutuality of engagement was associated with improved physics interest. In particular, students in groups in which each partner “participates equally in all aspects of the lab, including operating the apparatus and recording and analyzing data” (the equal participation variable), on average, reported an increase in their interest in physics. Given the positive outcomes associated with improved physics interest [124], this analysis suggests that instructors should ensure that students experience equal participation in all aspects of their lab work, and not be satisfied with students merely splitting their work 50/50.

The third part of the analysis focused on how equal participation predicted peer effect on self-efficacy controlling for E-CLASS self-efficacy and gender. We showed that men saw no added benefit when they participated equally in all the different aspects of the lab, including data collection and analysis, but that women did experience such a benefit. Given that women currently experience a variety of disadvantages in physics learning environments, including the masculine culture of physics [115] and stereotype threat [195], we suggest that ensuring equal participation in all aspects of lab work can help to improve equity in physics labs by elevating the self-efficacy of women without negatively affecting the self-efficacy of their male classmates.

These results suggest that student learning in introductory physics labs that includes a higher mutuality of engagement is likely to be accompanied by higher average levels of physics interest for all students, and by improved physics self-efficacy for women. In turn, development of interest and self-efficacy may lead to improved physics identity and other related short and long term professional outcomes, especially for women. For this reason, we suggest that instructors who wish to address inequities in their physics labs should seek to ensure that students are able to participate equally in all aspects of the lab work, including operating the apparatus and recording and analyzing data. Returning to the titular question, we conclude that a good lab partner is someone who will not just split up the work, but who will share equally in all aspects of learning in the introductory physics lab.

6.0 Group Diversity and Engagement

6.1 Introduction

Research has demonstrated the value of collaboration in helping students learn physics [82, 85, 206, 255, 267, 272] Collaboration gives students opportunities to learn from their peers; contribute to projects that are too large or complex to tackle alone; and share insights, approaches, and ideas. Collaboration is also an essential element in introductory physics lab courses [1, 222], in which students typically work in small groups to conduct experimental work.

A significant body of research has been conducted to address the issue of collaborative work with diverse teams [36, 40, 287]. Whether a diverse team of students is able to work productively together depends on a large number of factors, including the fraction of team members from historically underrepresented groups, the type of work being done, and the context of the study [91]. Past research has suggested that groups of three with two men and one woman in physics classes were dominated by the men [125, 126].

Other research has documented barriers faced by members of historically underrepresented groups in STEM disciplines such as physics when they collaborate with peers from overrepresented groups, including in labs [62, 74, 141, 236, 314] Examples of social interactions that may serve as barriers for members of underrepresented groups include microaggressions, having ideas ignored, or being talked over [24, 61, 74, 246] We expect that such barriers might decrease the effectiveness of collaboration in diverse teams by undermining group cohesion [122].

In order to develop an understanding of the relationship between group diversity and the effectiveness of collaborative work in an online introductory physics lab, we collected reports from a collaborative activity during the Fall 2020 and Spring 2021 semesters at a large research university in the USA. Each week, students worked in a randomly-assigned group of 4 via a Zoom breakout room to complete an activity that required collaboration. We extracted information about the level of engagement from the group reports ($n = 1342$).

In addition, we conducted interviews with students about their experiences engaging in this collaborative work ($n = 11$). Unlike in Fall 2020, in the Spring 2021 semester the instructor provided students with detailed grading rubrics, and so we compare engagement for different group compositions with and without rubrics. In this study, we seek to address the following research questions.

- RQ1: For students in the online introductory physics lab course, what was the relationship between the gender composition of their group and the degree of engagement they demonstrated in their reflective reports for collaborative activities?
- RQ2: What was the impact of providing students explicit grading rubrics on patterns of engagement for different group compositions?
- RQ3: What experiences did students describe in interviews that may illuminate the relationship between the gender composition of the group and the contributions of group members in shaping the group report?

6.2 Theoretical Framework

To address these questions, we adopt a communities of practice framework [51, 149, 244]. In this framework, we center our analysis on groups of learners as the "basic building block of a social learning system" [300]. Communities of practice consist of three dimensions: membership, practice and mission [301]. The membership of our community of practice is the group of 4 (or 3 if the number of students is not a multiple of 4) students who are randomly assigned to work together on the collaborative lab activity. The practice in our community of practice is the work they do together, in this case the collaborative group activity in the physics lab. If it is to be effective, the practice of the community should be aligned with the mission of the community work (i.e., the course goals), to strengthen students' understanding of physics concepts and experimental physics through the collaborative lab activities. In this study, we were interested in how the diversity in group composition was related to how well the group was able to collaborate productively and prepare a reflective report.

A key element of a community of practice is engagement [300]. Skinner and Belmont describe engagement, in an educational context, as “sustained behavioral involvement in activities” [273]. Inspired by efforts in education and social work [46, 193, 210] we operationalize engagement in our community of practice by measuring the relative length of the reflections sections of students’ reports. We hypothesize that groups with more engagement will be likely to generate more ideas, develop more-elaborate reasoning, or persist longer in documenting their discussions, all of which could lead to longer responses in the reflections. Likewise, in groups with lower engagement, or sustained involvement, we might expect cursory, short responses in the reflections, on average.

Response length on reflections is a straightforward and objective measure of engagement, allowing us to collect data from the hundreds of groups needed to make quantitative comparisons between different types of groups. We observed lab groups as they conducted the collaborative activity, and noted that groups that engaged more in discussions during the reflection also wrote longer responses. Therefore, since it was objective, meaningful, and possible to evaluate for hundreds of groups, we used collaborative reflections response length as a measure of group engagement.

We analyzed group engagement for groups with different gender compositions in the online lab course. We obtained anonymized gender information from institutional records in which students’ genders were recorded as male, female, or unknown. All of the students in our course identified as either male or female. Thus, while gender is a fluid and non-binary construct, for the sake of this analysis we rely on records that report all the students in our study as either men or women. However, our perspective in this study is that gender is something that students *do* [115, 302]. Based upon their gender identity, students may interact differently with other group members in the collaborative online lab.

6.3 Context

This study was conducted in an introductory physics lab course in the Fall 2020 and Spring 2021 semesters. Because of the COVID-19 pandemic, the course was taught entirely online. Students completed lab investigations and homework independently, using the IOLab device and materials from RealTime Physics [278]. Once per week, students joined a Zoom meeting in which they were randomly assigned to breakout rooms with 4 students (sometimes three if the total number of students was not a multiple of 4) in order to work on a collaborative activity. The lab sections were taught by graduate student teaching assistants (TAs). Random assignment was decided upon as a practice since TAs felt that assigning students to groups in Zoom was complicated and time-intensive. At our university, students take two semesters of lecture-based introductory physics courses and one semester of physics lab (that students can take concurrently with their second lecture-based course or after taking that course). Therefore, the lab focused on a wide range of physics concepts such as mechanics, circuits, and optics. The TAs visited the breakout rooms occasionally, and were available if students had questions or needed help.

A total of 245 students (139 women and 106 men) participated in the lab course in the fall, and 377 (233 women and 144 men) in the spring, with a small number withdrawing from the course each semester after the first two weeks. The majority of students in the lab course were enrolled in health science tracks and a smaller number of students were physical science (e.g., chemistry or physics) majors.

The collaborative activities were designed to provide students with opportunities to work together with peers, given that the rest of their work in the lab course was done individually. Specifically, the collaborative activities aimed to give students opportunities to pool and reflect on data collected in different ways; design and conduct a new experiment as a team; and engage in group reflection on topics related to the nature of science [7, 213], group-work, or physics concepts. The activities were written for the IOLab Lesson Player [8, 32, 181], with three stages corresponding to the three goals: integrating analyses, group experiment, and reflections. The reflections stage always asked students to have a discussion about a given topic, and then to write responses. Some reflections prompts are summarized in Table 6. It was the responses from the reflection stage that were used as data for this study.

During the Fall 2020 semester, students were provided with a generic grading rubric on the first page of the collaborative activity. The TAs graded student work generously, rarely assigning less than 90% for the collaborative work. For the Spring 2021 semester, specific grading rubrics were developed to support the TAs in more carefully grading each collaborative activity. In the first three weeks of the Spring 2021 semester, the average scores for the collaborative activity were 88%, 89%, and 82%. Following an outcry from students, the instructor decided to release the specific grading rubrics to the students at the start of the fourth week of the Spring 2021 semester. By the sixth week, the average score was 94%, and it stayed above 90% for the duration of the semester. We hypothesize that grading protocol, the availability of specific grading rubrics, and implicit messaging in the rubrics that student work would be graded partially for thoroughness might have caused a change in how students oriented with respect to their work and, accordingly, to the nature of their collaboration. Therefore, in the analysis that follows, we compare data from Fall 2020 with data from weeks 6-12 of Spring 2021, after students were provided with the grading rubrics.

6.4 Methodology and Results

6.4.1 Quantitative Analysis

In total, 1342 reports from 622 students working in groups of 3 or 4 were collected and analyzed. Not all students were present each week, but all students participated in the majority of the collaborative activities. Students who withdrew or dropped the class were excluded from the analysis, as were students who completed the activities individually (e.g., making up late work). We determined the composition of each of the 1342 groups according to the genders of its students, and arranged those groups on Fig. 3. For example, groups whose composition was “One Woman” were groups of three that had one woman and two men, and groups of four that had one woman and three men. The middle column in these graphs, “Two Each”, was only possible for groups of four.

To parameterize engagement, we extracted the students’ responses to the reflection questions from the third stage of the collaborative activities. We calculated the number of characters in each response. We decided to count the number of characters rather than the number of words because we wanted to acknowledge that responses using precise technical

Table 6: Examples of reflection prompts for the collaborative activities. Some prompts have been condensed for brevity.

- How does the IOLab device know its position and velocity?
- Why are predictions so useful and important in experimental sciences like physics?
- Why is it important to use theory to make predictions when doing science?
- (after a lab about Newton's third law) How important is it for passengers on buses to wear seatbelts?
- Why do you think doing a physics lab can be valuable for you as a student who is learning physics? What are the goals of the physics lab?
- (after a lab about electric circuits) What are three rules to describe voltage and current in electric circuits?
- Are "rules of thumb" useful in science? Explain.
- What does it mean for a measurement to be uncertain? Why is this important, and what are the implications for science?
- Setting aside the issue of grades, what aspects of this lab helped and hindered your learning of physics?

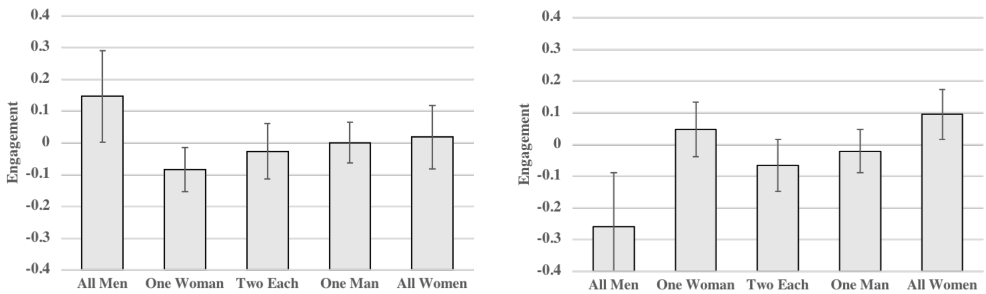

Figure 3: Engagement, as measured by the length of reflections, Z-scored, for five different group compositions, according to gender, for the Fall 2020 (left) and Spring 2021 (right) semesters. Error bars indicate standard error. In the Fall semester, groups with one woman demonstrated lower engagement. In the Spring semester, all groups with women demonstrated comparatively higher levels of engagement.

jargon might use fewer, but longer, words. To account for the likelihood that groups of 4 would write more than groups of 3, we divided the number of characters from each response by the number of students in the group. We tested this hypothesis by comparing average characters-per-person in groups of 3 and 4 and found that they were very similar, week-by-week. Next, we calculated a Z-score for each group using the mean and standard deviation for each week, i.e.: $Z_i = (x_i - \bar{x})/\sigma_x$. So, in a particular week, if a group of 4 students wrote a response that was 3080 characters long, we calculated a characters-per-person of 770 for that response. The average number of characters-per-person in week 5 was 600 and the standard deviation was 170, so we would record an engagement Z-score of +1.0 for the group in that week.

Finally, we compiled all of the weeks together, for the two semesters, and created the charts in Fig. 3. The charts show average engagement Z-scores and standard errors for different group compositions. The columns for each chart average to 0, weighted by the number of groups in the columns. For each chart, we calculated a one-way ANOVA [263] to determine the significance of differences in engagement between different group compositions. Given relatively small differences, we did not find statistical significance at $p < 0.05$ for groups with different gender composition. Although non-significant at $p < 0.05$, the data in Fig. 3 suggest that in the Fall 2020 semester, groups that were all men had higher levels of engagement and groups with only one woman had lowest engagement. In the Spring 2021 semester when students were provided detailed rubrics, the data suggest that groups that consisted of all men had lowest levels of engagement and all other group types had comparable levels of engagement.

6.4.2 Qualitative Methodology

In order to investigate the experiences of students that might have impacted different engagement levels in groups with varying composition, we conducted a series of 11 semi-structured interviews [228, 259] at the end of the Fall 2020 semester. Each interview was 60 minutes long, conducted via Zoom, and audio-recorded. The author conducted 5 interviews, and the author's supervisor conducted 6 of the interviews. We invited students to participate in the interviews, offered to pay $25 for each interview, and asked interested students to email us. 22 students expressed interest. We selected 6 women students for the author's supervisor to interview. Next, we randomly selected 6 students from the remaining 16 students for the author to interview. One of those randomly-selected students decided to withdraw from the study after the start of the interview. We ended up interviewing 9 women and 2 men.

We recognize that our positionalities are relevant to how we conducted the interviews and analyzed these qualitative data. We have removed the remainder of this paragraph, describing the positionalities of the author and his supervisor, to align with guidance about masking this manuscript for double-blind peer review. The author is a white man who was a graduate student at the time of the interviews, and whose perspective in conducting education research is influenced by a decade teaching high school physics. Students might have recognized him as a person who is related to the labs because he recorded videos for the lab course and wrote the collaborative activities. The author's supervisor is an Asian woman physics professor who has taught and conducted physics education research since 1995. Students would not have had classes with her previously, but might view her as a champion for gender-based inclusion in physics.

The interview followed a set of questions relating to the students' background and expectations about physics lab, their impressions about the educational effectiveness of the online lab-work, their opinions about learning online, and their experiences collaborating with classmates while doing online labs. We transcribed the interviews and conducted coding using NVivo. In the first cycle, we used concept coding [251] and identified 19 emergent codes, such as "Group Roles" and "Zoom Etiquette", to describe the concepts students described in their interviews. In the second cycle, we used the quantitative findings that groups with only one woman has less engagement as a starting-point to engage in elaborative coding [251]. In this second cycle, we read the transcripts for students' descriptions of gendered experiences in collaborative work. Finally, after identifying key passages in the interviews, the author and his supervisor discussed how they connected with our theoretical framework for this investigation and past research results.

6.4.3 Qualitative Data

Most of the women we interviewed described gender-based discrimination of one form or another, usually within the physics lab. Within the theme of gendered experiences, students' responses generally fit into two categories: being excluded, or stereotypes about behavior or performance. In our interviews, several women described incidents in which they were ignored, excluded from conversation, or talked over by the men in their group. One woman described being in a group with one other woman. The men in the group would exclude the other woman, ignoring her ideas or questions. "[She] was just talked over and [she] wasn't given as much consideration. Nobody would take the time and explain it to [her]... And they would just push through it. [She] would have to deal with it."

A challenge experienced by all students during this online course was dealing with group members who kept their cameras off and stayed muted so that it was hard to know if there was anyone there. One of our interview participants, a woman, explained how faceless critique from a man in her group was detrimental to their collaborative work. "There was this one group that had one guy [who] didn't say much... He only un-muted if he saw something wrong with us doing something." This man was not contributing to the collaborative work in a constructive way; rather, the only times he talked he instigated arguments and doubts about the procedure, as if he saw himself as superior to the other members of his group who were all women.

In another case, an interview participant described not being taken seriously by the men in the group until the TA came in and agreed with her. "I was the only girl and there were three or four other guys... I would say something which would be right... But other people would disagree with what I said. Which is fine, like, they can disagree. But I also felt like if it was one guy against me, all the other guys would just side with that guy even though they didn't even have a proper reason for doing it... Then the TA would come in and corroborate what I said. And then people would be like, okay, and then change it... I felt that a lot." Despite being correct, this woman needed an authority figure to agree with her before the other members of her group took her ideas seriously, and this was a frequent occurrence.

In a second category, some students described navigating a preconception that women tended to write more, and that they had higher expectations for the quality of their lab-work. One interview participant commented, "I noticed that the women students write longer descriptions than male students."

Another interview participant suggested that men she worked with were more likely to be content with imperfect work, commenting, "I think everyone that I've worked with wants to get a good grade. But I think sometimes when I do work with guys-only groups, if we don't understand something, they're like, 'oh well, what we've put is good enough."' She went on to compare this to her experience working in a group of three women: "When I worked with just two girls, they were like, no, we have to make sure this is right." Another interview participant agreed with the preconception that women were more likely to put in the effort to ensure their work was thorough. She told a story about a time when, in a group in which she was the only woman, she went back to correct the group work after they had completed it, but before it was turned in. "I was the one who had to send it to [male partner] so that they could upload it. And after they had ended the call and I was looking through it, I realized that I would not put a lot of the answers that they put, because a lot of the answers that they put were very skimpy, not very detailed, and didn't include a lot of the stuff that I would have told them to include. So I had to spend like an extra like 10-15 minutes trying to like fix the answers a little bit." The student elaborated, comparing her work with groups of all women with groups in which she was the only woman: "When I'm with girls I don't worry about that as much, but when I'm with guys, I do try to insert more and, be like, let's add this, let's add this. Because I don't feel like they necessarily will."

6.5 Discussion and Conclusion

In the quantitative portion of this study, we find suggestive, but statistically non-significant results, that suggest groups with only one woman had decreased levels of engagement in online group work. However, when explicit grading rubrics and a clear grading incentive were in place, mixed-gender groups demonstrated comparatively higher levels of engagement. We hypothesize that this result suggests that women, who may be more grade-conscious, are empowered to push their mixed-gender groups to write longer reports when it is clear that their grade depends on it.

In the qualitative portion of this study, we found some evidence suggesting that gender diversity in collaborative online lab work impacted students' level of engagement. For example, the women we interviewed described instances in which they, or other women they knew about, were excluded or talked over while attempting to engage in collaborative work. Our hypothesis is that if a group excluded or ignored the ideas or contributions of one group member, this may have directly reduced the level of engagement we measured and the reflection response would be shorter.

Furthermore, based upon interview data, we hypothesize that awareness of stereotypes about their contributions or expectations may have caused some of the women in these groups to feel as if they were walking on eggshells while collaborating. The women, hyper-aware of their difference within the masculine culture of physics, may have carefully weighed their participation in the collaboration. This hesitancy may have reduced the engagement in the group, resulting in women contributing fewer ideas, being less likely to engage in discussion, and writing less when they were the only woman in a group. As a result, for groups in which women were in the minority and stereotypes were present, the engagement of the group as a whole may have been reduced. In other words, we suggest that cohesion of the collaboration between students [122] was reduced, in part, by the presence of stereotype threat, whereby the fear of conforming to a stereotype about one's social group can influence one's behavior [197, 284, 285] However, in Spring 2021, when students were provided with explicit rubrics for how their work would be graded and were provided a grade-based incentive to submit their best work, groups with women no longer had a lower level of engagement demonstrated by reflective report length.

Although further studies are needed, these findings hint that instructors can improve equitable participation of women in their lab classes by either not forming small groups in which there is only a single woman [126], or ensuring that students are provided with clear performance expectations via explicit rubrics to assess their work.

7.0 Lessons from Transforming an Honors Physics Lab

Labs designed for physics majors are essential for preparing the next generation of physicists. These labs should provide physics students with opportunities to learn to think like a physicist in a lab context and teach them essential skills that will be useful for both academic and non-academic careers. However, in typical labs encountered by physics majors, including labs beyond the first year, one typically finds complex, expensive equipment that bear little resemblance to what would be used in an academic research lab or on a job site. While undergraduate research experiences may provide a bridge between undergraduate classes and research labs,[146] they often pose the same difficulty of either requiring or at least preferring skills and familiarity with equipment that isn't typically found in lab courses.[189]

Thus, there is a need to transform instructional labs in order to better equip physics majors with skills they can use in undergraduate research experiences, graduate school research, or non-academic jobs. The creative and purposeful use of computers and/or low-cost electronics is a popular and successful approach for retooling physics labs,[11, 25, 34, 70, 103, 109, 211, 275, 281] and is part of the strategy we adopted to transform honors physics lab with three new lab modules at our institution.

At the University of Pittsburgh, first-year physical science students (including physics majors) take a separate 2-credit lab course after their first introductory college physics course rather than taking introductory labs as part of, or in parallel with, their first introductory physics classes. Thus, the first physics lab that physics majors take is either the regular introductory lab in the spring semester of their first year of studies or the honors lab, discussed here, that is offered in the fall semester of their second year of studies. This honors lab requires a certain minimum grade in the introductory physics lecture course and has a reputation of requiring intense work, with two three-hour lab sessions weekly. Enrollment is typically between 10 and 16 students. Given the advanced preparation of many of its students, the topics of investigation, and the scheduling during the second year of studies, the honors lab may best be compared with Beyond the First Year (BFY) labs at other institutions even though it is the first college physics lab taken by the students enrolled in it. Approximately 30% of our physics majors take this honors lab, while the remaining 70% take a regular, non-honors, lab which also enrolls chemistry, engineering, and other majors. 87% of the students in the honors lab are physics majors, while only 20% of students in the regular, non-honors, lab choose to major in physics.

The process of transforming a physics lab has been well-explicated in the physics education literature.[137, 322] The first step was to identify goals by consulting with stakeholders such as faculty members and referring to documents such as the AAPT lab guidelines.[1] Second, we focused on designing new lab work, procedures, and apparatus consistent with the feedback obtained. Finally, we evaluated the transformation, which involved three new Arduino-based lab modules as well as additional features such as an 'above and beyond' task associated with each lab.

Two overarching goals led our efforts: helping students learn to think like physicists, and teaching essential undergraduate-level research techniques to students. These goals led to two guiding principles in the design of the three new modules in our transformation. First, that equipment be research-grade, used across multiple experiments, and not merely a 'black box'. Second, that students be able to troubleshoot equipment, explore variables and parameters in the experiment, and work collaboratively. To address the first set of principles, an Arduino-based digital test instrument was developed and deployed. All software developed for this transformation is open source and available online,[89] and additional specifications for how the system is set up are available on request. To address the second set of principles, changes were made to the way the course was run, including the adoption of an 'above and beyond' requirement for lab work.

The perspective of the cognitive apprenticeship model,[56] which was useful in this context of transforming the lab, proposes that learning is effective when the criteria of good performance are modeled explicitly and then students are provided coaching and scaffolding support to learn important skills before they can successfully practice those skills independently. We believe that scaffolding new skills is essential for students in lab courses. Both troubleshooting and the 'above and beyond' work will be productive for students only if they are provided with the guidance, scaffolding and support they need in order to be successful in the new, transformed lab. To evaluate the impact of the three new Arduino-based modules and 'above and beyond' task associated with the transformed lab, we analyzed student work, evaluated student attitudes toward experimental science using the E-CLASS survey, and conducted observations and interviews of students enrolled in the lab.

7.1 New Lab Modules

The simplicity, low-cost, ubiquity, and capabilities of the Arduino Due make it an effective instrument for the physics lab.[33] While previous work [102, 119, 142, 177] has primarily focused on using the Arduino microcontroller boards for exciting one-off investigations, we report here on a flexible and powerful system that brings research-grade electronics to some of experiments students perform in the second-year honors teaching lab. In combination with a simple shield (a custom-built board which interfaces with the Arduino Due, see Fig. 4), a breadboard, and a computer, the Arduino-based system is able to replicate the capabilities of a variety of traditional lab equipment. Data are transferred from the Arduino to a computer, where open-source software [89] performs analysis to replicate the functions of an oscilloscope, a synthesizer, a lock-in amplifier, a spectrum analyzer, and a network analyzer. The focus is on helping students learn to deeply understand and debug issues related to the software of one piece of equipment, rather than poring over the details of many complicated devices.[65, 66] Encouraging and teaching troubleshooting skills was a major goal of the lab transformation.

While the lab seeks to help students discern value in both computer-based data acquisition systems and electronic test equipment such as multimeters and oscilloscopes (both of which are used frequently in the second half of the course), there are several advantages to helping students learn to use the Arduino-based instrument first, rather than using a collection of separate lock-in amplifiers, spectrum analyzers, network analyzers, and other expensive test equipment for lab work. The breadboard ecosystem allows students to use inexpensive, easily-replaced components, and the Arduino-based system itself is inexpensive in comparison with lock-in amplifiers and network analyzers that are available on the market today.

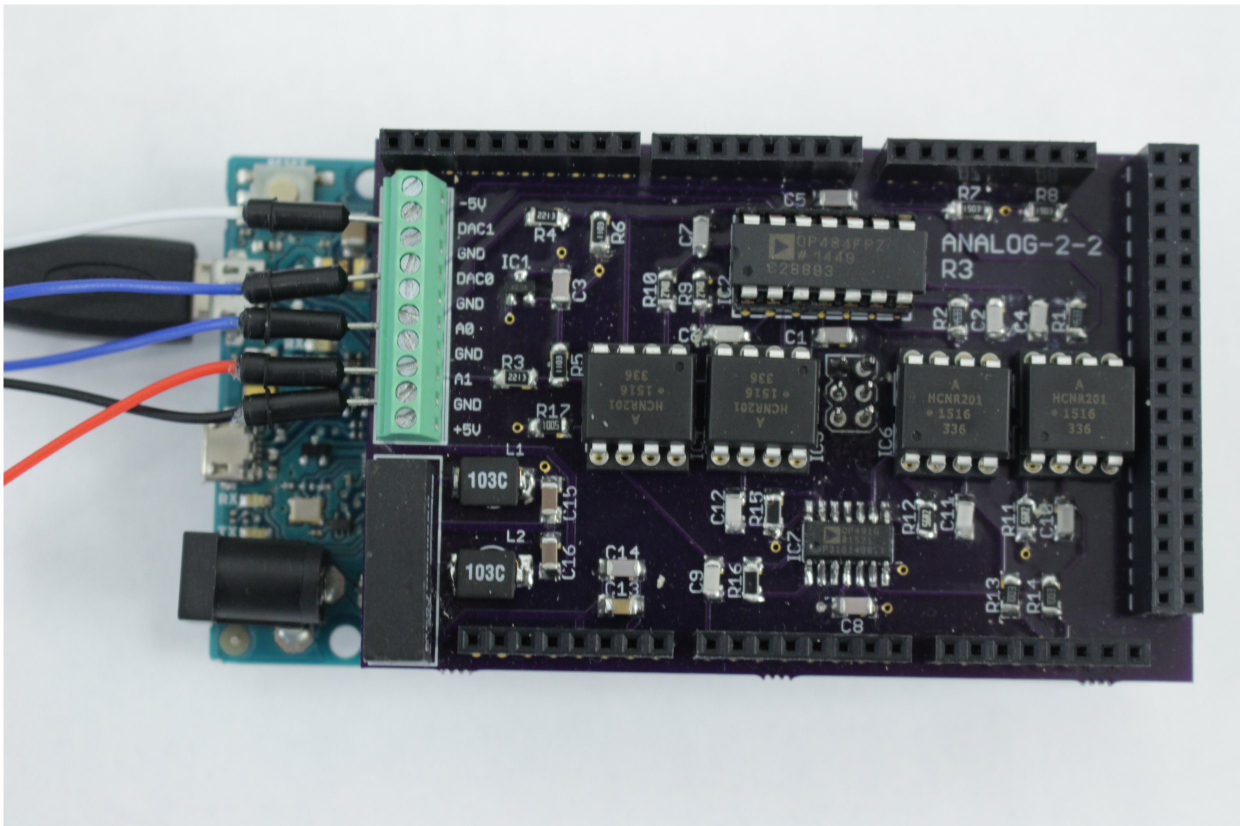

Figure 4: The Arduino Due and custom-built shield used for input and output and analog electrical signals in the honors physics lab.

Students are guided and encouraged to view the Arduino-based instrumentation as a tool that they can dig into and understand: they are taught some basic Python in the first weeks of the course and encouraged to explore and edit the data analysis code that runs on the computer, and to do the same with the data acquisition code that runs on the Arduino. Furthermore, students typically co-enroll, or take in a subsequent semester, physics courses in scientific programming and digital electronics which use Python and the Arduino, respectively, so that – at least in principle – students should be able to understand and work with the Arduino-based instrumentation at a deeper level by the time they are ready to begin undergraduate work in a faculty research lab. A major goal of introducing these new lab modules that use the Arduino-based instrumentation was to encourage students to move beyond thinking of their data acquisition system as a black box.

The Arduino Due itself is built around a powerful microcontroller of a type that is commonly used in academic and industrial research labs (an ARM Cortex-M3, with a 280 kHz sample rate and 12-bit analog input/output bit depth), so that expertise with Arduino systems can be ported directly into research-grade lab work.

Students, in general, responded positively to the introduction of the Arduino in the three new modules of this honors lab course, expressing enthusiasm about their lab work and indicating excitement about coming to the lab course. Some students reported that they were able to take their lab skills to research labs and immediately start working on research. More than 50% of students, and 100% of students in the most recent offering of the course, elected to perform one or more Arduino-based experiments when given the choice (see Table 7.)

Below, we describe the three new experiments that were designed to use the Arduino-based instrument. They supplant experiments that relied on equipment for which parts are no longer available, or for which traditional equipment was not up to the task of adequately performing standard analyses. The experiments are also designed so that students are slowly introduced to new capabilities of their instrumentation and the features of the experimental apparatus, which allows them to learn to think systematically about how their instruments may be misbehaving and troubleshoot more effectively. All three new experiments were carefully written to meet our overarching goals of helping students learn to think like physicists

Table 7: Experiments available to students in the honors physics lab course.

- Test Measurements*
- Single Photon Interference
- Noise Fundamentals
- RLC Circuits*
- Electron Diffraction
- Nuclear Magnetic Resonance
- Acoustical Resonance*
- Photoelectric Effect
- Chaotic Circuit
- Acoustical Gas Thermometer
- Black-Body Radiation
- Muon Lifetime
- Electron Spin Resonance
- Microwave Optics
- Radiation Detection

and teaching students techniques useful for experimental research.

7.1.1 Test Measurements

After a four-week introduction to lab procedures and software, including an introduction to Python programming and data analysis techniques for physics, students work through this first lab to learn how to use a variety of test equipment. This equipment includes an oscilloscope, function generator, multimeter, spectrum analyzer, lock-in amplifier, and network analyzer, all of which run on the Arduino-based instrument. After some introductory measurements, students pass signals from a conventional or Arduino-based function generator through an RC circuit and qualitatively measure the amplitude and phase response at a given input frequency using an oscilloscope. Then they use a spectrum analyzer and lock-in amplifier to examine the response more quantitatively. Next, the network analyzer function is used to measure the amplitude and phase response over a wide range of frequencies and to show how RC circuits can be used as low/high-pass filters. Finally, students are given freedom to explore the Fourier decomposition of a square wave or repeat the analysis using an RL circuit.

This lab is used to introduce the Arduino-based instruments to students. The use of a conceptually-simple circuit makes it easier to scaffold student understanding of both the functioning of the new instruments as well as the signal processing that is being performed. The Arduino-based lock-in amplifier, spectrum analyzer, and network analyzer make it possible to have all students complete this lab at the same time, without needing to share costly equipment. This lab introduces students to a number of important fundamental principles about electronics that they will revisit in their advanced-level electronics lab course, including equipment such as breadboards, instrumentation, and methods for troubleshooting. After this experiment, students rotate among a selection of two-week experiments, completing four, for which there may be only one set of apparatus available. These experiments are listed in Table 7 (but may not all be operational at any given time). The following two labs are included in the rotation.

7.1.2 RLC Circuits

In this two-week lab, students first examine the transient response in an RLC circuit using the synthesizer and oscilloscope. After that, they drive the circuit with a sine wave and investigate the steady-state response using the oscilloscope, spectrum analyzer, and lock-in amplifier. They also use the network analyzer to see how this circuit behaves around resonance. Some emphasis is placed on measuring and understanding the phase response of the circuit and comparing the response to that of a harmonic oscillator. The goal is to combine elements of electronics (which will, as with the Test Measurements lab, be built upon in the advanced electronics lab course) with the interesting physics of resonance.

While the Test Measurements lab takes students through a scaffolded, step-by-step procedure, this experiment is more open-ended. After introducing the theory and illustrating the circuit to build, the instructions give students a few specific issues to investigate, while giving students plenty of freedom about which tools to use, the specific parameters to use, and so forth.

7.1.3 Acoustic Resonance

In this two-week lab, students examine acoustic resonant modes in a wooden box using a speaker and microphone. Students use the oscilloscope, and eventually the lock-in amplifier and network analyzer, to measure the frequencies of modes, as well as the phase response. Finally, the students move the microphone around the box in order to map out a resonant mode in physical space. This serves as a useful contrast to the RLC circuits lab, as it brings the concept of resonance into physical space for students to explore. As with the RLC Circuits experiment, this lab is also quite open-ended, with students given plenty of scope to explore and address questions they find interesting on the way through the procedure. While the RLC circuit has a simple, single resonance, the acoustical cavity has much increased complexity due to the presence of many resonances with varying spatial distributions.

7.2 Student Learning

Alongside the new instrumentation, several elements of the course were changed with the goals of the transformation in mind. Lab handouts were rewritten from a highly-structured format to a more informal discussion of topics to investigate. The grading scheme was changed to emphasize the importance of maintaining a useful lab notebook.[283] The change to Arduino-based instruments softened the black box nature of instrumentation that is typical in physics labs, which necessitated more time spent upfront helping students learn how the Arduino-based instrument works.

A benefit from this approach to understanding the instrumentation is that it makes troubleshooting a central, and repeatedly-emphasized, element of the lab course. The instructor of the course motivated the transformation by focusing on the importance of troubleshooting and how it can help students learn to think like physicists. Students need guidance in order to learn explicitly how to troubleshoot their apparatus, and scaffolding needs to be provided for learning troubleshooting techniques.[66, 83] Troubleshooting was explicitly taught in two ways. First, in the Test Measurements module, students were stepped through the experiment in such a way that they could learn how to test the ways in which individual electronic components affected an electrical signal. They used probes to make measurements at several points as a signal passed a circuit, and were asked to explain how and why the signal changed at each point. A common and important form of troubleshooting in this lab was tracing signals through electric circuits on the breadboard in order to diagnose circuit wiring difficulties. Second, lab instructors made a point of helping students learn to diagnose and address issues with their experiments in a supportive way throughout the course. One way that instructors provided this assistance was by suggesting specific measurements that students could make in order to produce results that would be helpful in diagnosing the issue they were encountering.

Another change was the requirement that 20% of students' lab reports discuss explorations 'above and beyond' the scope of the work assigned in the lab handout. Students each wrote a complete traditional lab report for their final experiment, which was graded using a rubric that focused on the clarity of the description of the lab work and the correctness of the analysis. The 'above and beyond' requirement aligns well with the goal of getting students to think like physicists, as this independent exploration is an exemplary opportunity for them to develop their curiosity, their skills in designing and conducting experiments, and their physics identities.[44, 124, 155, 156] For example, a student working on an interference experiment did a calculation to show that there was (usually) only one photon present in the device at a time. Another example was when a student noticed an unexpected behavior in one of the graphs, and followed up with some insightful analysis of how the Arduino might have a non-negligible internal resistance or inductance.

Using a generative coding scheme [228] developed and validated by the author and his collaborators, an analysis of 28 lab reports from one semester of the transformed course in which students completed more than one lab report found a large variety of 'above and beyond' work being done by students, as shown in Fig. 5. Some students chose to investigate and write about the theoretical side of their results, relying on their textbook and internet research to demonstrate a deeper understanding of the underlying theory. Some reports showed students following up on a result that caught their eye by attempting to understand what might have caused that result. Other approaches included taking additional data to extend the range of their investigation, performing a deeper dive into the error analysis, writing about potential improvements that could be made to the apparatus, and providing a narrative-style description of one part of the experiment that did not work out as expected. Most impressive were the lab reports in which students made a substantial, meaningful, and well-explained extension to their lab work. In these cases, the students truly went 'above and beyond', demonstrating independent investigation skills and aptitudes that show they are well-prepared for a research setting.

However, the limited amount of scaffolding provided for the 'above and beyond' work made it difficult for a sizable number of students to excel in this aspect of the lab course. These students, whose work shows up in the approximately 30% of Fig. 5 categorized as "no 'Above and Beyond' seen", may have needed more guidance and feedback, and may have benefited from seeing examples of what this type of work is, or could have benefited from some additional support regarding the type of work that would be reasonable or acceptable. In the most recent offering of the lab course, examples of past 'above and beyond' work were

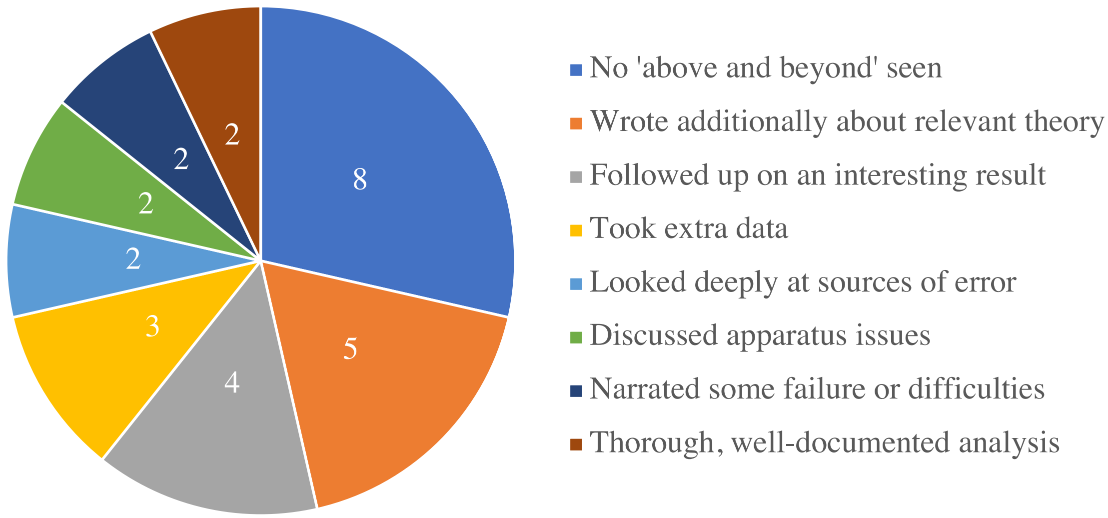

Figure 5: Types of 'above and beyond' work done by students in a transformed honors physics lab.

provided, and this was viewed as quite useful by the students.

In order to evaluate how student attitudes toward experimental physics have changed as a result of the transformation, the E-CLASS survey was given to students at the beginning (pre) and end (post) of the semester. E-CLASS scores indicate the extent to which participants hold expert-like views about experimental physics, including strategies, habits of mind, and attitudes as measured through 30 items with a 5-point Likert scale.[308] In keeping with precedent, we express scores from -1 (novice-like) to $+1$ (expert-like) and focus on students' own responses, and not how they believe an expert would respond, which the E-CLASS survey also captures. Fig. 6(a) indicates scores from a nationally-representative sample [311] and Fig. 6(b) shows scores from our university. For the honors physics lab, our results suggest a larger decrease in E-CLASS scores the year prior to the transformation, 2015, (effect-size given by Cohen's $d = 0.67$, t-test $p = 0.08$ [55]) and a smaller decrease during post-transformation years, ($d = 0.10$, $p = 0.31$). Although the number of students in these two samples is small, and so the results are not statistically significant, the fact that the effect size for the decrease given by Cohen's d goes from 0.67 to 0.10 is encouraging. The E-CLASS scores for the honors lab are similar to those for BFY courses in the national sample ($d = 0.02, p = 0.10$; compare BFY in Fig. 6(a) with Honors in Fig.6(b)). The difference as measured by Cohen's d in this case is "small".[55] Analysis of the national sample in Fig. 6(a) conducted by Wilcox and Lewandowski indicated that students in regular, traditional labs typically see a decrease in E-CLASS scores, while labs that include open-ended work (like our transformed labs) produce scores that are unchanged or increase slightly between pre and post[311].

Moreover, the selective nature of the honors lab is apparent in Fig. 6, as E-CLASS pre scores for our honors lab are consistently higher than pre scores in both first-year (FY) labs from the national sample [311] (t-test $p < 0.001$) as well as our regular, non-honors, lab ($p < 0.001$). In addition to being selective and potentially attracting students with high levels of prior preparation, the majority of the students who enroll in the honors lab are male physics majors. Most female physics majors choose to enroll in the regular, non-honors, introductory lab, in which physics students account for 20% of the enrollment. However, even though students in the honors lab may be considered a privileged group overall compared with the regular introductory lab, that doesn't mean they all have sufficient prior experience, e.g., with experimental techniques, and don't need to be supported in their learning in the lab context.

7.3 Student Experiences

The physics lab can be a culturally-rich, low-stakes environment for students to develop useful research skills and stimulate their interest in physics. Formative lab experiences, like learning to use the Arduino, can help some students to develop an interest in physics and come to see themselves as physicists or scientists. From the beginning of the transformation, students generally responded positively. Most students who were introduced to the Arduino boards quickly became familiar with the platform, and were soon able to put the device to use in a variety of creative ways. Some students explicitly expressed enthusiasm for the transformation and indicated that they looked forward to coming to the lab. Other students

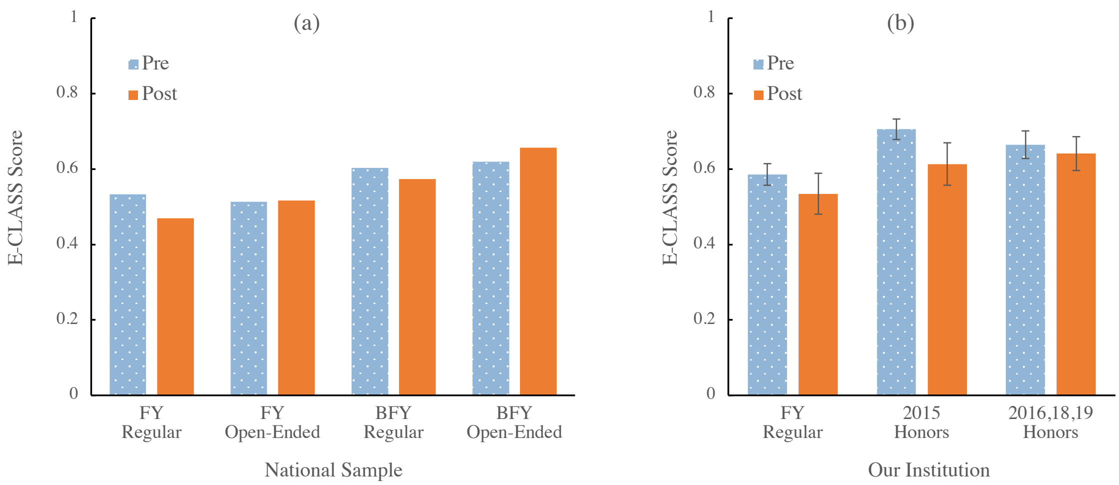

Figure 6: Pre and post E-CLASS scores for a national sample of students separated into First-Year (FY) and Beyond First Year (BFY) labs, and whether students primarily do regular, traditional lab work or open-ended experiments. Also includes scores from our institution include a regular, non-honors, introductory lab that is taken by physical science majors, including physics students who do not take the honors lab, the honors lab before transformation, and the honors lab for three offerings after the transformation.

subsequently sought out research opportunities and reported that they were able to immediately begin working productively in a research lab using skills and understandings they had developed in class. However, for some students who may not have received appropriate support, lab experiences can negatively impact their physics identity trajectory.[74] In order to understand the impact of this lab transformation on student experiences, we conducted 12 hours of observations in the lab and interviewed 7 students who took the transformed lab.

The lab environment generally seemed positive both before and after the transformation, with some male students cracking jokes (often involving experiments and troubleshooting) and good relationships between instructors and students. Students were free to work alone or in pairs, and nearly all chose to work with a partner. During years with an odd number of students, this typically meant that one student would work alone. After the transformation, the instructor repeatedly emphasized the importance of learning to troubleshoot apparatus: being patient, learning from mistakes, and being deliberate and methodical in learning how to find and diagnose potential issues from unexpected results. While troubleshooting is indeed a skill that we need to help students learn, it can be frustrating and overwhelming if a student does not have a partner or they don't feel comfortable asking for help, since students may not have the self-confidence and prior skills to effectively carry out the troubleshooting task without appropriate guidance and scaffolding support. In particular, one difficulty with our initial approach was that it put the onus on students to ask the instructor for assistance if they were struggling.

For students who end up working alone, and for those with lower levels of self-efficacy, it sometimes meant that students didn't feel comfortable asking for help. Based on discussions with students, we found that this affected two students in the first years of the lab transformation. A male student who did not have a partner because he was the 'odd one out' in his year described the lab as very difficult and felt very negatively about his experiences. He struggled to complete the experimental work and write up a lab report by himself. This is not surprising since the amount of lab work was designed to be shared by two students. A female student from a different year of the lab transformation, who described this class as "totally awesome" in principle, explained how being the only woman in the lab course meant that her male classmates “kind of paired off”, leaving her to work alone. As an unsupported solo student, she reported that the course “was just really hard... there isn’t that much guidance, and that can make it really difficult, especially because I was doing it without a partner.” She described struggling with the equipment and having difficulty managing the troubleshooting while working alone, saying, “If I don’t know what I’m doing, it’s hard for me to say what help I need because I don’t know where I’m going wrong in this.” She added, “You start off and you have the oscilloscopes and you have to hook up a bunch of stuff with cables, which I had never done before, and I didn’t know what all these cables were and where you’re supposed to get them, and it was a ton of just learning how to do measurements.” Frustrated by the struggle of working alone on a lab designed for two students, her experiences led her to question whether she belonged in physics labs, saying “I just didn’t feel like I was supposed to be there at a certain point.”

While both the male and the female student found working alone to be difficult, being left alone to do all the work had a more detrimental effect on the female student. Perhaps because of a lower sense of belonging and a lower level of self-efficacy,[157, 205] the female student dropped out of the course, enrolling in the regular introductory lab instead, while the male student in the same situation was frustrated by being left alone but managed to persist and complete the class.

We note that, since it may be more difficult for underrepresented students such as women to find a partner to work with in labs such as these, it is important that instructors be deliberate and thoughtful in ensuring that students are not left to work alone if they are not adequately supported to succeed while doing so. After identifying this challenge, we took steps to ensure that subsequent offerings of the honors lab would allow students the opportunity to work in groups of 3 if necessary, that the instructor would form groups if anyone was being left out, and that the instructor would take initiative to visit students working on labs in order to provide more support to students who need it, whether they ask for it explicitly or not.

In the most recent offering of the course, the female students who enrolled in the course found partners and no students dropped out because of needing to work alone. Moreover, two female students who worked together in the most recent offering of the transformed lab described their experience in the class as a very positive one. They appreciated that the lab instructor, aware of the challenges faced by students in this lab, frequently checked in with them and was willing to answer their questions quietly in a corner of the room where they wouldn't be overheard by other students. For students with a lower sense of belonging, it can be particularly valuable to be able to ask questions without fear of being judged by their peers. Similarly, a female student who worked with a male partner described the lab as being a good course, challenging but fair.

7.4 Discussion and Summary

Based on our experiences, we offer the following advice for instructors of similar labs who wish to ensure their students are being given equal learning opportunities and not being overly burdened when they have students work in groups:

1. Ensure all students have an opportunity to benefit from working with a partner,[267] even if they might initially prefer to work alone.
2. Assign students to groups if needed, respecting that students may benefit from choosing their own partners when they work in pairs.[14] For groups of more than two, be careful to avoid marooning underrepresented students (e.g., a woman working with two men) and attend to the roles students assume in their group work.[125, 74] As the female student who dropped the transformed lab after being left to work alone in the earlier implementation of the lab explained, "Making the teachers assign partners, as uncool as that is, would definitely help".
3. Rotating groups occasionally (e.g., every few weeks) can help students break out of bad work habits and learn to work with different partners.[125]
4. Check in with students frequently. Seek to make it normal to ask questions and seek guidance without fear of being judged, and take advantage of the opportunity to provide guidance to students out of earshot of their peers where they can feel more comfortable (this may be particularly beneficial for underrepresented students).

Labs that target physics majors have the potential to provide students the opportunity to develop experimental skills and expertise with research-grade equipment and help them learn to think like a physicist. The Arduino-based instrument described here is a flexible and powerful device that helps to meet this opportunity. Meanwhile, changes in the structure of the course such as supporting the development of troubleshooting skills, scaffolding 'above and beyond' work, and adopting open-ended skills-based work could help students to develop essential scientific skills for future research in academic and non-academic settings.[49, 167]

While the spectrum of types of 'above and beyond' work shown in Fig. 5 is impressive, we are in the process of both improving the overall quality of this type of work and helping students who do not undertake 'above and beyond' work for all their lab reports. We may think of this challenge as seeking to balance the innovative aspect of student work with the need for students to practice skills to improve their efficiency in the work they do in the lab.[257] In other words, we need to ensure that the struggle students experience in the lab is productive [160] and that students are not frustrated with the open-ended nature of the troubleshooting and 'above and beyond' tasks. Moving forward, we plan to provide improved support and guidance for students as they learn to troubleshoot their apparatus, such as by playing out simulated problems in a coordinated way.

Past research suggests that active learning (of which lab work is an example) can contribute to decreasing the performance 'gap' between overrepresented and underrepresented groups of students.[191] However, further investigations suggest that it is the implementation of active learning that is the critical factor in determining whether active learning is able to close these performance gaps. In particular, one study [161] showed that while all students learned more in evidence-based active learning classes, the performance gap between men and women increased from pre to post test. This is the main reason we focused explicitly on the experiences of individual students in our evaluation of this transformed honors lab via individual interviews and lab observations. More generally, our findings support the claim that failing to attend to the needs of traditionally-underrepresented groups of students in the lab risks perpetuating inequities in physics.[74] For the honors lab, which is taken by many of our physics majors, this first college experimental physics experience is critical for students who may already struggle to see themselves as physicists.[44] Therefore, we conducted individual interviews with a subset of students in order to highlight the importance of attending to the needs of underrepresented students even if there are very few such students. The struggles of these types of students will not be captured by aggregate data such as the E-CLASS scores in Fig. 6.

The transformation of a second-year honors lab curriculum offers an exciting opportunity to increase the relevance, accessibility, and quality of an essential physics learning experience. It is important for such transformations to be done in a way that ensures that students are adequately supported and carefully accounts for the complete processes of anticipated skill development. By doing this, new lab courses such as the one introduced here can provide a venue for all students to develop both positive physics identity and valuable research-ready lab skills.

8.0 Lessons from Transforming Introductory Labs

Recent research on introductory physics labs suggests that students are neither learning physics concepts nor developing expert-like attitudes toward experimental science [134, 313]. One criticism leveled at introductory physics labs is their "cookbook" nature, whereby students follow a series of directions in a lab manual, producing results without understanding the underlying physics concepts or engaging with the scientific process at anything other than a superficial level [135]. Notable efforts to move beyond the cookbook approach have focused on building inquiry-centered learning environments [3, 8, 35, 96, 214].

This work focuses on a calculus-based introductory lab, offered at our university as a separate, 2-credit, course for chemistry and physics majors. Enrollment during the semester of investigation was 30% female. Students attend a weekly 1-hour lecture in which the instructor gives an overview of the relevant physics topics to be encountered in the lab that week, and a 3-hour lab session where they work with a partner at a computer-equipped lab bench. The labs are run by graduate student teaching assistants (TAs).

Our first effort to reform this course was the introduction of 6 electricity and magnetism labs from the inquiry-based Real-Time Physics curriculum [278]. These replaced cookbook labs during weeks 5-10 of the 12-week sequence, and served as a trial before securing funding for apparatus to switch to a full implementation. Students completed worksheets from the Real-Time Physics lab guide, and also did pre-lab exercises and post-lab homework from the guide. Neither TAs nor students received any special training for this style of lab, nor were efforts made to motivate the switch or get "buy-in".

Here, we present results from a series of reflexive ethnographic-style observations [39] and pre/post attitudinal surveys. The observations shed light on students' behaviors and the social dynamics in the lab while the attitudinal survey helps us to identify students' beliefs about the nature of lab-work and their lab experiences. Taken together, these results help to illuminate what students are thinking and doing in their lab classes, and thus guide further reform efforts.

8.1 Methodology

Ethnography: The first half of this work is based on approximately 100 hours of observations spread over the same semester as the survey administration. These observations were performed using an ethnographic protocol adapted from the field of cultural anthropology [39]. Given the potentially subjective nature of such work, the observer must be reflexive: that is, adopt "an approach to participant observation that recognizes that we are a part of the world we study" [38].

Consequently, it is essential for the observer to strike a balance between involvement in the culture being observed, on one hand, and affective detachment from it, on the other [39]. The observer becomes a natural and accepted figure, while still retaining the ability to make observations that are as unbiased and as revealing as possible. Conclusions are reached by collaboratively evaluating the observer's field-notes and impressions while taking into account the observer's background and the context for the observations.

In the labs, the observer (D.D., a graduate student) introduced himself as a researcher interested in monitoring and improving the lab experience, and positioned himself as a friendly but taciturn fixture of the lab-room. He sometimes sought students' opinions on the work they were doing, and occasionally answered student questions or stepped in when students were at risk of doing something dangerous.

Mostly, however, the observer sat at the side of the room: watching, listening, and recording notes. He was careful to avoid interfering with TA-student interactions or with the students' lab-work. The observer's experience with inquiry-based instruction at the high school level meant that he was readily able to discern the cookbook labs' inability to engage students in sense-making. On the other hand, as a white male, it took him longer to start recognizing aspects of psychosocial interactions such as microaggressions.

Several times through the semester, the observer and the collaborators performed a reflection activity designed to consolidate observations and identify relevant threads in ethnographic research [194]. Some threads, such as gender dynamics and student-TA interactions, prompted focused attention to aspects of the lab in future observations. At the end of the semester, a meta-reflection was performed on the observations and reflections to weave these

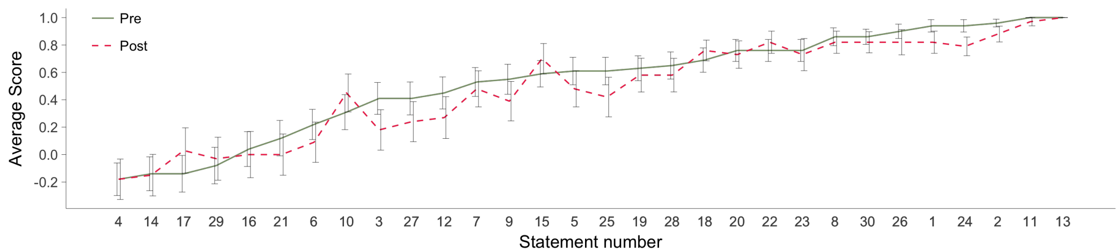

Figure 7: Average E-CLASS pre and post scores, with the statements ordered according to ascending pre-instruction score, for transformed physics labs.

threads into a report about the student experience in the lab.

E-CLASS: The quantitative portion of this work is based on an E-CLASS [308] survey administered in the second and second-to-last weeks of the lab course. The survey was distributed in the last 10 minutes of the lecture, and students were asked to indicate their responses on bubble sheets. The survey was anonymous: no demographic information was collected, nor was any incentive provided for completion.

As a research-validated instrument, the E-CLASS is designed to probe student expectations and epistemologies related to lab-work and the role of experiments in science [308]. The survey asks students to respond to statements such as ”When doing an experiment, I try to understand how the experimental setup works.” Although the original study asks students to respond in additional ways, in order to keep the survey to a reasonable length, we asked students only to respond to the 30 statements from their own perspective.

Responses are indicated on a 5-point Likert scale, and compared with the expert-like response. The ”strongly agree” and ”agree” responses are aggregated, as are the ”strongly disagree” and ”disagree”, and accorded points such that each question is valued at +1 if the student's response is expert-like, 0 if neutral, and -1 if the student's response is novice-like. Averaged over all students, the result is a score from -1 (novice-like) to +1 (expert-like) for each of the 30 statements.

Themes: After the ethnographic protocol produced a set of relevant pedagogical themes in the lab classes, a team of 8 PER researchers was asked to classify each of the 30 E-CLASS statements according to those themes. The researchers "agreed" on their classification if at least 7 of the 8 researchers identified a statement with the same theme, and no more than 3 of the researchers also classified the statement with a second theme. A Fleiss' kappa test was used to assess the inter-rater reliability between the 8 raters [97].

8.2 Results and Discussion

The synthesis of ethnographic observations identified a number of key themes in student lab experiences common to cookbook-style and inquiry-based labs. First, a recurring theme was the degree to which students demonstrated agency in their lab-work. The lab manuals simplified the work and thinking expected of students, and both the preceding lecture and TA support further narrowed the scope of the learner's agency. Consequently, students were rarely required to make decisions about how to collect, process, or present their data, and often struggled when such decision-making was required.

Second, as the semester progressed, we noticed a decrease in some students' willingness to undertake lab tasks, attempt explanations of complex concepts, or take initiative in completing lab work. This decreased engagement was oftentimes gendered: for example, a female student who is increasingly withdrawn as a male colleague takes over the apparatus. Recent work from our group reported that the self efficacy of female (but not male) students decreased significantly during physics classes at this level [201]. Thus, given how self efficacy can inform learner engagement, we determined that the self efficacy of female and underrepresented minority students should be an important point of reference.

Third, we saw a number of students misunderstanding the nature of scientific knowledge-generation in experimental physics. For example, some espoused the belief that the purpose of experiments was simply to confirm known results. Since this was explicitly the purpose of much of their cookbook-style lab-work, it is possible that the lab was reinforcing undesirable beliefs about the nature of science. This agrees with recent findings in related work [143, 144].

Fourth, we identified a spectrum of fundamental lab skills, with some students failing to correctly read fundamental measuring devices like calipers or multimeters.

These four themes (learner agency, self efficacy, nature of science, lab skills) may be important dimensions for further reform effort. Therefore, the researchers sought to determine whether these themes could be identified in the E-CLASS survey. If we could identify statements that correspond to particular themes, scores on those statements could be used to guide and evaluate reform efforts. In total, 10 of the 30 statements met these criteria: four statements that the researchers associated with self efficacy, and six that were associated with nature of science. No statements were associated with the other two themes to this stringent level of agreement.

The inter-rater reliability on the 10 statements for which the researchers found agreement gave $\kappa = 0.68$ (substantial agreement [175]), indicating that this reduced categorization scheme is a meaningful one. Thus, it is reasonable to use these 10 statements from the E-CLASS survey to track the extent to which our students' self-efficacy and understanding of the nature of science are being impacted by the lab course.

8.2.1 E-CLASS Results

A total of 49 valid responses were obtained from students in the second week of the lab course, and 33 valid responses in the second-to-last week of the course. Three responses were discarded because the student penciled in the same response for each statement. This represents a majority of the students in the lab class. The initial enrollment was 56, decreasing slightly to 48 by the end of the course.

We compared the results from our implementation of the E-CLASS with the national norms established in Ref. [308]. Averaging over all the responses to all the statements, we find that our pre and post scores are each indistinguishable from their national norms. Given that the post condition reflects the impact of 4 weeks of cookbook-style labs and 6 weeks of the inquiry-based investigations, this result indicates that the overall effect of this admixture of learning tasks was not different from "business as usual" cookbook labs.

We also compared E-CLASS pre and post scores. The overall effect is a decrease in expert-like responses, with the average score decreasing from 0.53 to 0.48 (on a scale from -1 to +1). These results are similar to the national norm [308]. Item-level responses are presented in Fig. 7.

Table 8: E-CLASS statements identified as relevant to Self Efficacy.

2	If I wanted to, I think I could be good at doing research.
9	When I approach a new piece of lab equipment, I feel confident I can learn how to use it well enough for my purposes.
13	If I try hard enough I can succeed at doing physics experiments.
24	Nearly all students are capable of doing a physics experiment if they work at it.

8.2.2 Self Efficacy

The four statements identified as belonging to the theme of self efficacy are listed in Table 8. On statements 2, 13, and 24, our students exceeded the national norms on the pre-test. Statement 9 is narrowly contextualized to the use of lab equipment, and has a lower pre-test score than the national norm. The average score on these items decreased from 0.86 to 0.77, in line with the national norm [308].

One possible reason for this decrease, suggested by our observations, may be the prevalence of microaggressions in lab social interactions. Some examples we observed included male students increasingly taking over control of the experimental apparatus from their female partners, students of color being snubbed by peers while choosing their lab partners, and TAs responding differently to male and female students.

These observations point to the importance of TA preparation that includes equity and anti-bias training in setting up and managing the lab as a sociocultural environment. Moreover, in evaluating further reforms, we will look at responses to these four statements as a source of information about the degree to which the lab may be differentially affecting the self efficacy of female and underrepresented minority students.

Table 9: E-CLASS statements identified as relevant to the Nature of Science.

16	The primary purpose of doing a physics experiment is to confirm previously known results.
22	If I am communicating results from an experiment, my main goal is to make conclusions based on my data using scientific reasoning.
23	When I am doing an experiment, I try to make predictions to see if my results are reasonable.
26	It is helpful to understand the assumptions that go into making predictions.
28	I do not expect doing an experiment to help my understanding of physics.
30	Physics experiments contribute to the growth of scientific knowledge.

8.2.3 Nature of Science

Six statements were identified as being related to the nature of science (Table 9). The students scored well on these statements (> 0.50), with the exception of statement 16, which is about lab-work confirming previously-known results. Since much of the lab-work drew on theory the students had already seen multiple times, this novice-like response on statement 16 actually corresponds to their experience of experimental physics in this course.

Our results show the average score on these items decreased slightly from 0.66 to 0.63. However, since our lab course is designed to help students learn about the role of experimentation in the nature of science, we might hope that scores for these statements would increase. Even though the scores are mostly expert-like, the importance of the nature of science in an experimental physics course means this is nonetheless a theme to be addressed.

Our ethnographic observations suggest that one source of this novice-like thinking may be that students entered the lab excited to do experiments, but were disappointed to find that their work was routinized and simplified. They rarely confronted phenomena, theory, or experiments that are not already outlined in a standard textbook, and typically found themselves asking questions such as, "What does the lab manual tell us to do next?" rather than doing sense-making and asking "How can we understand this more meaningfully?"

Thus, we plan to modify the labs and implement tasks that more-closely model understanding of the nature of science we wish students to adopt during the lab. We also plan to introduce activities that will help students make connections between the experimental physics done in the lab and the model of scientific knowledge production we wish to promote.

Table 10: Low-scoring E-CLASS statements associated with inadequate skill development in cookbook-style labs.

14	When doing an experiment I usually think up my own questions to investigate.
17	When I encounter difficulties in the lab, my first step is to ask an expert, like the instructor.
21	I am usually able to complete an experiment without understanding the equations and physics ideas that describe the system I am investigating.
29	If I don't have clear directions for analyzing data, I am not sure how to choose an appropriate analysis method.

8.2.4 Impact of Cookbook-Style Labs

Our observations also suggested that students rarely spent time investigating phenomena that weren't explicitly mentioned in their lab manuals. Likewise, we saw that students often had difficulty troubleshooting their apparatus. Similarly, it was rare to see students make connections between the equations of the underlying theory, on one hand, and the resulting graphs and calculations, on the other. In the case of the cookbook labs, this may have been because the procedure was simplified so much that such connections were already made for them in the lab manual. These observations suggest that cookbook-style labs are not adequately helping students to learn the skills indicated in the AAPT recommendations for labs [1]. As shown in Table 10, four of the six lowest-scoring E-CLASS statements reflect these skills and attitudes.

8.2.5 Impact of Inquiry-Based Labs

The Real-Time Physics inquiry-based sequence is, in some ways, the opposite of a cookbook lab: it focuses on concept development, and interactions with experimental apparatus are mostly unstructured. Real-Time Physics labs intersperse instructions with questions related to the physics theory, which we observed to promote meaningful and engaging discussion about physics concepts: students were much more likely to engage in conversation about physics concepts with their peers during the six Real-Time Physics labs.

Nonetheless, we cannot separate the impact of this inquiry-based approach from cookbook-style labs, as the E-CLASS post scores do not differ from the national norms. This may be because our implementation of Real-Time Physics was for only half of the course, and that we didn't plan for specialized TA training, student and TA "buy in", or the targeted development of some specific lab skills. Our results emphasize the difficulty of implementing an inquiry-based approach to lab-work.

8.3 Conclusions and Future Plans

Initial steps were taken to transition an introductory lab course from a cookbook-style experience toward one driven by inquiry and meaningful learning. Our E-CLASS survey data suggests that the piecewise-adopted inquiry-based curriculum was not successful in achieving these goals. Our ethnographic observations strengthen this claim, and suggest that the causes may be related to microaggressions and social dynamics, counterproductive messaging about the nature of science, and other issues related to the structure of the labs.

We have identified three directions for future growth. First, we have begun to develop a robust TA training module to ensure that student inquiry is being supported effectively and fairly. Second, we have started creating small supplemental learning activities so students can explicitly learn about the nature of science and develop lab skills (e.g., how to make quantitative comparisons). Third, we offered a full sequence of Real-Time Physics labs in the following academic year. Meanwhile, the E-CLASS survey will allow us to monitor the impact of our efforts on the self-efficacy of female and underrepresented minority students, on students' understanding of the nature of science, and on our success in inculcating expert-like attitudes and lab skills.

9.0 Attitudes Toward Experimental Physics in Inquiry-Based Labs

In recent years, there has been renewed interest in introductory college physics lab courses [136], leading to the development of a variety of formats and pedagogical approaches for introductory physics labs in a similar way to development and analysis of approaches for lecture-based classes [158, 161]. In part, this interest has been driven by research that suggests students do not learn physics concepts in traditional, highly-structured physics labs [134, 274]. In addition to addressing the need for students to learn physics concepts, another important course goal that motivates some lab transformations is the inclusion of scientific inquiry in lab experiences [1, 104]. By inquiry, we refer to experimental work designed to probe and illuminate the validity and functioning of scientific models or hypotheses [10, 171].

Corresponding to the various goals of lab transformations, a variety of evaluation tools have been developed and employed. Concept inventories such as the Force Concept Inventory (FCI) [130], the Force and Motion conceptual Evaluation (FMCE) [290], and the Mechanics baseline Test (MBT) [129] have been used to determine how much labs have helped students learn physics concepts. Assessments like Physics Lab Inventory of Critical Thinking (PLIC) [297] and Physics Measurement Questionnaire (PMQ) [296] have been used to measure students' lab skills, such as critical thinking and uncertainty analysis. Surveys such as the Colorado Learning Attitudes About Science Survey for Experimental Physics (E-CLASS) [308, 313] and the Maryland Physics Expectations Survey (MPEX) [240] have been used to evaluate changes in students' expectations about physics and attitudes toward aspects of experimental physics. It is noteworthy that traditional physics labs see a decrease in students scores on the E-CLASS [308] and other attitudinal surveys, suggesting that un-transformed lab instruction might push students to adopt less-expert-like views and attitudes about experimental science.

One approach to transforming lab courses is to integrate lab and lecture, creating a learning environment that allows students to engage in scientific inquiry while also building up their understanding of physics concepts. Since inquiry is integrated into the course and classrooms are rearranged to promote collaborative work, such approaches typically do not have separate lab courses. Some examples of this approach include the Investigative Science Learning Environment (ISLE) [96], Modeling Instruction [35], Physics by Inquiry [214], Student-Centered Active Learning Environment with Upside-down Pedagogies (SCALE-UP) [26], Studio Physics [179], or Technology-Enabled Active Learning (TEAL) [27]. This approach has been associated with improved conceptual learning for students and unchanged (i.e.: not decreasing) attitudes toward experimental physics [309]. However, due to financial, scheduling, or other constraints, it may not be possible for all institutions to switch from a lab-and-lecture introductory physics sequence to an integrated course.

For introductory lab courses that are not integrated, many transformed lab courses have adopted conceptual inquiry-based curricula. Conceptual inquiry-based lab curricula focus on providing students opportunities to conduct experiments as a way to develop their understanding of physics concepts [104]. Evaluations of conceptual inquiry-based labs have reported on student conceptual understanding using the FCI [45, 226, 243, 294], MBT [226, 243], and FMCE [278]. However, little work has been done to evaluate how inquiry-based labs may affect students' attitudes toward experimental science. Student attitudes toward, and understanding of, experimental can be an important learning outcome for lab courses, and can also moderate the effectiveness of lab work on conceptual learning. One study reported that E-CLASS scores increased in an inquiry-based lab but, as the authors argue, the context of that study makes it difficult to extrapolate [264]. A large-scale study by Wilcox and Lewandowski shows that students' attitudes toward experimental science decrease in "guided" introductory physics labs and do not decrease in "open-ended" labs [311]. Another study using the same data shows that students' attitudes decrease in introductory physics labs in which the purpose is to "reinforce physics concepts" but increase very slightly in labs in which the purpose is to "develop lab skills" [312]. However, since many conceptual inquiry-based labs do not fall neatly into the guided/open-ended/concepts/skills categorization scheme, it may be difficult to infer definite conclusions from this work about the impact of conceptual inquiry-based lab curricula on students' attitudes toward experimental physics.

Thus, we seek to address the research question: How are students' attitudes toward experimental physics impacted by a conceptual inquiry-based curriculum, as measured by the

E-CLASS?. In addition, we seek to shed some light on two additional questions: What is the impact on students' attitudes of adding questions to a conceptual inquiry-based curriculum that ask students to reflect on the nature of their experimental work? And what is the impact on students' conceptual understanding in this conceptual inquiry-based introductory physics lab of different lab curricula?

9.1 Materials and Methods

At many institutions, the introductory physics sequence is offered as a two-semester series of lectures with associated labs. In contrast, at our institution, a large public research-intensive university in the USA, the introductory physics lab is a single course that is not associated with the introductory lecture courses, although there is overlap in the physics concepts covered. The lab requires that students are simultaneously enrolled in, or have completed, the second semester of the introductory physics lecture course. Thus, all of the students who are enrolled in the lab have already completed the first half of the introductory physics sequence, and have completed their studies of the topic of mechanics.

In our labs, students collaborated in groups of 2 (or 3, if necessary), to conduct their experimental work. They wrote a group lab report during the three-hour lab period. Understanding the nature of student collaboration in introductory labs is beyond the scope of our analysis in this chapter, but has been reported elsewhere [74, 81].

9.1.1 Three Lab Curricula

In our study, students completed three different lab curricula. The first of these, which we will refer to as the traditional curriculum, is a set of highly-structured lab exercises that were written by the author's collaborator [50]. In the traditional curriculum, students followed step-by-step instructions as they set up apparatus, made measurements, and performed calculations. Measurements and calculations were submitted digitally using LON-CAPA [172], which performed some checks to ensure measured values were reasonable and calculated values were correct. Since they were given step-by-step instructions, students were able to operate sophisticated equipment, including oscilloscopes (measuring the speed of sound in a cardboard tube), teltron tubes (determining q/m for an electron), and spectrometers (spectra from discharge tubes). The lab manual provided a review of the relevant physics each week, but students often found themselves ritualistically following the instructions from the lab manual rather than thinking about physics concepts. The traditional curriculum may be akin to introductory physics lab instruction as it has been practiced in many colleges and universities over the past century [31].

The second curriculum represents our best effort to faithfully implement 12 inquiry-based labs from the RealTime Physics curriculum [276, 277, 279, 280]. The design of these labs was informed by physics education research and, indeed, RealTime Physics has been shown to boost student understanding of physics concepts [278]. In our implementation of RealTime Physics labs, students completed a pre-lab activity individually, conducted the lab-work in pairs, and then completed a post-lab homework assignment individually. Each of these three tasks was graded. The lab-work called for students to develop multiple representations of physics concepts (e.g., ticker tape-style dot diagrams, velocity-time graphs, and written descriptions of motion). The lab-work frequently used a cycle whereby students made a prediction about the outcome of a simple phenomenon (e.g., what are the forces when two carts collide?), and then conducted the phenomenon to check their prediction (e.g., collide two carts with force probes on a dynamics track). The RealTime Physics labs are structured, in the sense that students complete a series of tasks, including stating predictions, conducting small experiments, and answering questions that connect experimental results to physics concepts. Students used computer-based data collection with Vernier equipment, including dynamics tracks, sensors, and optics benches. In our implementation, students completed 6 mechanics labs, 3 electric circuits and electromagnetism labs, and 3 optics labs, outlined in Table 11.

The third curriculum represents our efforts to adapt RealTime Physics to the needs of our students and our course goals, which were to improve student conceptual understanding of physics and to help students learn skills and ways of thinking relevant to experimental physics. For this third curriculum, we kept the pre-lab and post-lab assignments from Real-

Table 11: The 12 labs included in the RealTime Physics and RealTime Physics + Reflection curricula.

Week	Lab
1	Changing Motion
2	Force and Motion
3	Combining Forces
4	Newton's Third Law and Conservation of Momentum
5	Two-Dimensional Motion
6	Conservation of Energy
7	DC Circuits
8	Capacitors and RC Circuits
9	Magnetism and Electromagnetism
10	Reflection and Refraction
11	Geometrical Optics
12	Waves of Light

Time Physics. We changed the lab-work that students performed in only one significant way, by adding reflection questions [159] to the lab-work. These questions called for students to reflect on the nature of science and ways of thinking in experimental science, and are inspired by the American Association of Physics Teachers Recommendations for the Undergraduate Physics Laboratory Curriculum [1] and recent scholarship about reflection and critical and scientific thinking in physics labs [84, 138, 144, 169, 248]. These reflection questions asked students to think individually, confer with their lab partner, and then write a group response, therefore requiring both collaboration and metacognition. We refer to this third curriculum as RealTime Physics + Reflections.

Five of the first six such reflections are presented in table 12. For brevity, surrounding text that unpacked and contextualized the questions has been removed. In pairs, students would write paragraph-long answers to these questions, which would be graded using a rubric. Alongside each reflection question, we also present an E-CLASS item that is closely aligned with the reflection question. During the fifth, and after the sixth week of labs, the reflection questions tackled topics such as the usefulness of simulations, the importance of varying only one variable at a time, and the process of theories becoming accepted in science. Since these topics are not closely aligned with E-CLASS items, they are not presented here.

9.1.2 Students

At our university, we offer two sets of physics courses: algebra-based and calculus-based. Most students in the algebra-based physics courses are health science majors (e.g., pre-medical students), while the calculus-based physics courses mostly enroll engineering and physical science students. These two streams, algebra-based and calculus-based, also have different physics lab courses. Thus, an algebra-based lab is made up mostly of health science majors. However, since engineering students are not required to take the introductory physics lab, the calculus-based lab course is made up mostly of physical science majors. Nonetheless, the lab curriculum is indistinguishable between the two streams, so that the only difference between the algebra-stream and calculus-stream labs is the pool of students enrolled in them. This research was carried out in accordance with the principles outlined in University

Table 12: Reflection questions from five of the first six labs from the RealTime Physics + Reflections curriculum.

Week	Question	E-CLASS item
1	Why are predictions so useful and important in experimental sciences like physics?	When I am doing an experiment, I try to make predictions to see if my results are reasonable.
2	Why is it important to think about sources of systematic error when doing physics experiments?	When doing a physics experiment, I don't think much about sources of systematic error.
3	Why is it important for scientists to use uncertainties when they analyze and share their work?	Calculating uncertainties usually helps me understand my results better.
4	Why is it important that physics students have opportunities to come up with their own experiments to investigate?	When doing an experiment I usually think up my own questions to investigate.
6	Some people believe that the purpose of doing a physics lab is simply to verify facts about physics that you already know... What is the goal of the physics lab?	The primary purpose of doing a physics experiment is to confirm previously known results.

of Pittsburgh Institutional Review Board (IRB) ethical policy with approval number/ID IRB: PRO15070212.

During the four semesters that we collected data, we adopted RealTime Physics in a piecewise fashion across the algebra- and calculus-stream labs. For the first two semesters, the algebra-stream labs used the traditional lab curriculum while the calculus-stream labs used the RealTime Physics curriculum. During the third semester, all labs used the RealTime Physics curriculum. During the fourth semester, all labs used the RealTime Physics + Reflections curriculum.

9.1.3 Other Potentially Relevant Factors

During the final semester in which we collected data, our university switched to emergency remote instruction due to the COVID-19 pandemic. As a result, the last five weeks of lab were completed asynchronously and individually, using simulations, rather than in pairs in-person. The pre-lab and homework were completed as normal, and 97% of students were able to complete all the lab-work remotely. Remote instruction at our university began March 23, and the E-CLASS survey was completed by students April 2-10. In the survey, we asked students about their attitudes toward the physics experiments they had conducted in the lab. Therefore, since students were exposed to a relatively short window of simulation-based asynchronous labs, and were asked to respond to the E-CLASS survey based on their in-lab experiences (which they had mostly completed), we do not believe that student responses during this fourth semester are substantially different because of the switch to emergency remote instruction.

The introductory physics labs at our university are run by graduate student teaching assistants (TAs). The TAs take a course on physics pedagogy during their first year in graduate school [204], and receive professional development related to their work supporting student learning in the labs during weekly lab TA meetings. This professional development focused on helping TAs support inquiry learning and inclusion in the labs, and is described elsewhere [75]. During the four semesters that data was collected, the professional development offered to lab TAs evolved only slightly. For example, during the first semester the

TAs wrote reflections during one lab TA meeting, whereas in subsequent semesters they discussed their reflections verbally. Thus, we do not believe that slight evolutions in the lab TA professional development would have impacted student survey responses.

Based on our experience with introductory physics labs at our university and the research literature, we have identified three additional factors, beyond those described previously, that may account for variation in how much students learn in the course. First, there are somewhat different pools of students in the fall and spring semesters for this lab course (similar to 'on-sequence' and 'off-sequence' enrollment [5]). Students in the spring semester are slightly more likely to be pursuing a pre-health science academic track (e.g., pre-medicine), are slightly more likely to be concurrently enrolled in the second semester of the physics lecture course while taking the lab, and have marginally higher high school grade point average (GPA), SAT Math scores, and grades from the first semester of physics. Second, we note that students who have stronger academic preparation are sometimes able to engage more productively in the lab [261]. Third, we note gender differences in how students approach and conduct their lab-work [61]. This suggests that, in attempting to understand the impact of our different curricula on student attitudes toward experimental physics, it will be important to account for differences in students' gender, academic preparation, and the semester in which they are enrolled.

9.1.4 Model and Mathematics

We propose a linear model [108, 288] to predict students' end-of-semester ($PostScore$) E-CLASS scores based on which version of the curriculum they experienced ($Inquiry$, $Reflection$). To account for the possibility of other significant factors impacting our results, and to minimize the potential impact of omitted variable bias [298], we also include variables for students' start-of-semester E-CLASS scores, the course (either algebra-stream or calculus-stream), the semester (either fall or spring), the student's high school GPA, and the student's gender. The model predicts coefficients (β), which may be interpreted as the predicted change in $PostScore$ associated with the factors, with all other factors held constant. Mathematically, this model takes the following form:

$$\begin{aligned} PostScore_i = \; & \beta_0 + \beta_1 \, Inquiry_i + \beta_2 \, Reflection_i \\ & + \beta_3 \, PreScore_i + \beta_4 \, Course_i + \beta_5 \, Semester_i \\ & + \beta_6 \, GPA_i + \beta_7 \, Gender_i + \epsilon_i \end{aligned} \tag{2}$$

PostScore and *PreScore* are student scores from the E-CLASS, a validated "expectations and epistemology" [321, 308] survey that aims to measure the extent to which students hold expert-like attitudes toward experimental physics. A total of 30 items are answered on a 5-point Likert scale, with +1 for each expert-like response (e.g., responding "strongly agree" or "agree" to an item for which the expert response is agreement), 0 for each neutral response, and -1 for each novice-like response (e.g., responding "strongly disagree" or "disagree" to an item for which the expert response is agreement [308]. The theoretical range of scores on E-CLASS is from -30 to +30. The validation of E-CLASS found an average score of 15.8 on the pre and 14.4 on the post for a national (USA) sample [308]. In interpreting the models in the results section, the coefficient of *PreScore* indicates the correlation between *PreScore* and *PostScore*. For example, a coefficient of 0.73 means that each point a student scores on the pre predicts 0.73 points on the post, on average, with all other variables held constant.

Inquiry is an indicator variable, taking on a value of 0 if the student was enrolled in a lab with the traditional curriculum and 1 if the student was enrolled in a lab that used the RealTime Physics curriculum (either with or without the additional reflection questions). *Reflection* is an indicator variable that takes a value of 1 if the student was enrolled in a lab that used the RealTime Physics + Reflections curriculum, and 0 otherwise. For both these variables, the coefficients indicate the number of additional points students earn on *PostScore* because they are enrolled in a lab that uses this curriculum, all else held constant.

Course is an indicator variable that takes a value of 0 for algebra-stream students and 1 for calculus-stream students. *Semester* is an indicator variable that takes a value of 0 for students enrolled in the fall semester and 1 for students enrolled in the spring semester. For these variables also, the coefficients indicate the number of additional points students are predicted to get on the post score because they are enrolled in a calculus-stream or spring semester class, all else held constant.

GPA is the student's high school grade point average, retrieved from university records through an IRB-approved process that anonymizes student data to preserve privacy. The intent of this variable is to provide an approximate accounting for the student's academic preparation. As a variable, *GPA* has been standardized to have a mean of 0 and a standard deviation of 1. Therefore, the coefficient of GPA indicates the number of additional points a student scores on *PostScore* corresponding to an increase in *GPA* by 1 standard deviation.

Gender is a binary variable that represents the student's gender, with women assigned 0 and men assigned 1. The coefficient indicates the number of additional points on *PostScore* for men, on average, with everything else held constant. While we acknowledge that gender is fluid and non-binary, we note that gender is nonetheless an important factor that can affect students' experiences in the lab and so we seek to account for it as best as we are able. In this case, we retrieved gender data from university records where it was stored as ”male”, ”female”, or ”other/unknown,” and excluded from our analysis the 4 or 5 (depending on the analysis below) students whose gender was stored as ”other/unknown.”

As a secondary analysis, we consider the impact of the three curricula on only the five E-CLASS items from Table 12 that are closely aligned with the reflection questions. We use Eq. 2, but note that the *PreScore* and *PostScore* in this analysis will have a theoretical range from -5 to +5.

In addition to the above analyses, we also look at the impact of the above factors on students' conceptual learning. For this, we use the Force Concept Inventory (FCI) [130]. While the only other study of RealTime Physics cited above uses the FMCE [278], most of the other work on inquiry-based introductory physics labs has used the FCI. Furthermore, our department has used the FCI in lecture courses for many years, and have typically found increases of 5-6 points as students complete the first semester of physics. The FCI consists of 30 items, each of which is either correct or incorrect, resulting in a score between 0 and +30. To evaluate student conceptual learning, we use Eq. 2, but with the FCI scores used as *PreScore* and *PostScore* intead of E-CLASS scores.

We calculated linear models using R Studio. All models were checked to ensure that errors were normally distributed and heteroscedastic. No statistically significant interactions were found in the models. Parsimonious models were developed by iteratively removing non-significant ($p < 0.05$) factors until all remaining factors are significant, with similar results. We present both complete models and parsimonious models, but note that in all cases the complete models maximized the Akaike Information Criterion [108] and are preferred for reporting results.

9.2 Results

We collected data over four semesters, from fall 2018 to spring 2020. Students filled out surveys during the lab sessions, and were given a small grade incentive for doing so. Students completed the E-CLASS pre on the first day of lab and the post on the second-to-last day. The FCI pre was also done on the first day of lab, and the post was done in week 7, after the mechanics labs were finished. Survey responses that were incomplete, that lacked a correct student identification number, or that had patterns suggesting the student did not provide honest answers (e.g., if all "A"s were selected) were removed. Pre and post responses were matched, and only matched data is presented here. We thus analyzed E-CLASS data from 701 students and FCI data from 569 students.

Table 13 provides panel data about survey respondents according to their course, semester, and curriculum. We have fewer responses from the calculus-stream labs because these labs have lower enrollment (engineering students, who make up the majority of students in our calculus-stream physics lecture sequence, do not take the physics lab at our university).

Table 14 shows results from our linear model for post E-CLASS score. Aside from the pre E-CLASS score, we find two statistically significant factors that predict post E-CLASS scores. $Inquiry$ is not a statistically significant predictor, indicating that students' E-CLASS scores are not different whether the traditional or RealTime Physics curriculum was used. However, $Reflection$ is a statistically significant predictor of post E-CLASS score. Thus, adding reflection questions to the RealTime Physics curriculum is associated with an increase of 3.80 points on the E-CLASS post score, all else being equal.

In the the models on Table 14, we find that students' gender is a statistically significant

Table 13: Number of responses for E-CLASS and FCI for the three lab curricula.

	E-CLASS	FCI
algebra, fall	254	134
algebra, spring	392	348
calculus, fall	25	38
calculus, spring	30	49
traditional	208	73
RealTime Physics	175	196
RealTime Physics + Reflections	318	300
total	701	569

Table 14: Coefficients from complete (first) and parsimonious (second) models for E-CLASS PostScore.

Factor	β	SE	p	β	SE	p
Inquiry	0.51	(0.81)	0.53			
Reflection	3.80	(1.07)	< 0.001	3.69	(0.49)	< 0.001
intercept	-0.15	(0.79)	0.85	-0.05	(0.61)	0.93
PreScore	0.73	(0.04)	< 0.001	0.74	(0.04)	< 0.001
Course	1.33	(0.93)	0.15			
Semester	-0.58	(0.89)	0.52			
GPA	0.06	(0.25)	0.82			
Gender	1.46	(0.52)	0.005	1.49	(0.51)	0.004
R^2		0.46			0.46	

predictor of post E-CLASS score. Men scored 1.46 points higher on the post E-CLASS than women, independent of their pre E-CLASS, high school GPA, and other factors related to the nature of their lab. The final statistically significant factor in Table 14 is $PreScore$, which significantly predicts $PostScore$.

We may also compare mean scores in order to interpret these results. We find that students in labs that used either the traditional or RealTime Physics curricula (without reflection questions) averaged a pre E-CLASS score of 14.3 and a post E-CLASS score of 11.4. Meanwhile, students in labs that used the RealTime Physics + Reflections curriculum averaged pre E-CLASS scores of 16.4 and post E-CLASS scores of 16.3 as well. To visualize these results, E-CLASS scores from algebra-stream students are presented semester-by-semester in Fig. 8.

In a second analysis, based on results presented in Table 15, we find that the RealTime Physics curriculum predicts a 0.33 point increase in students' scores on a 5-item subset of E-CLASS items related to the reflection questions. The RealTime Physics + Reflections curriculum predicts an additional 1.04 point increase in this score, for a total advantage of 1.47 points compared with the Traditional curriculum. As with overall E-CLASS scores, the difference comes from an averted decrease rather than an increase for the RealTime Physics + Reflections curriculum, compared to the traditional curriculum. Mean scores on the 5-item subset decreased from a pre of 1.20 to a post of 0.12 with the Traditional curriculum, but were unchanged, remaining at 1.68, with the RealTime Physics + Reflections curriculum.

In a follow-up to the main research question, we performed a similar analysis on students' post FCI scores. These model results are presented in Table 16. Once controlled for course, semester, preparation, and gender, we find that students in labs that used the RealTime Physics inquiry-based curriculum (with or without reflection questions) were predicted to have a post FCI score that was 2.60 points higher. In terms of mean FCI scores, this reflects an increase from 16.7 to 17.7, on average, for students enrolled in labs that used the RealTime Physics conceptual inquiry-based curriculum, compared with a decrease from 15.5 to 14.8, on average, for students enrolled in labs that used the traditional curriculum.

Students enrolled in the calculus-stream physics labs had post FCI scores that were 1.89 points higher than students in algebra-stream physics labs, controlled for pre FCI score. This

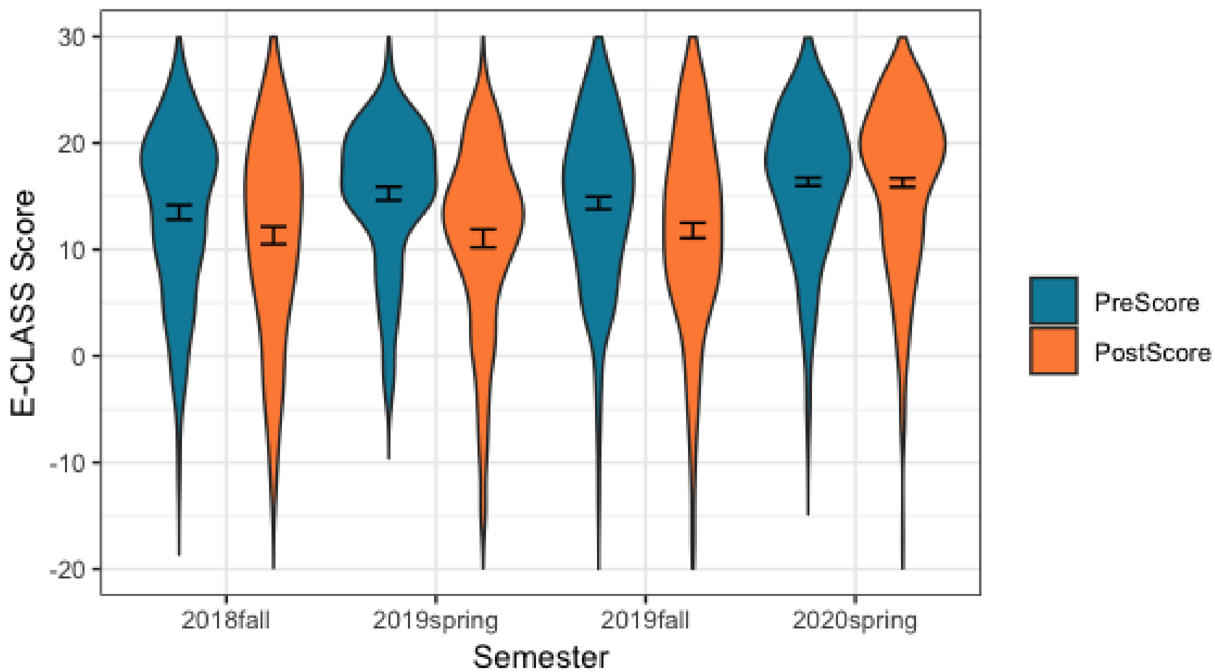

Figure 8: Violin plot of matched pre/post E-CLASS scores for algebra-stream students by semester.

Table 15: Coefficients from complete (first) and parsimonious (second) models for E-CLASS Reflection Question PostScore.

Factor	β	SE	p	β	SE	p
Inquiry	0.33	(0.24)	0.17	0.41	(0.20)	0.04
Reflection	1.04	(0.35)	0.001	0.92	(0.18)	< 0.001
intercept	-0.60	(0.20)	0.003	-0.67	(0.15)	< 0.001
PreScore	0.48	(0.04)	< 0.001	0.48	(0.04)	< 0.001
Course	0.17	(0.28)	0.54			
Semester	-0.13	(0.27)	0.64			
GPA	0.06	(0.07)	0.43			
Gender	0.54	(0.16)	< 0.001	0.53	(0.15)	< 0.001
R^2		0.28			0.28	

Table 16: Coefficients from complete (first) and parsimonious (second) models for FCI PostScore.

Factor	β	SE	p	β	SE	p
Inquiry	2.60	(1.29)	0.04	2.46	(0.69)	< 0.001
Reflection	-1.38	(1.15)	0.23	-1.23	(0.47)	0.01
intercept	3.25	(1.32)	0.008	3.23	(0.75)	< 0.001
PreScore	0.74	(0.04)	< 0.001	0.75	(0.03)	< 0.001
Course	1.89	(0.69)	0.006	1.85	(0.63)	0.004
Semester	0.12	(1.14)	0.91			
GPA	0.16	(0.21)	0.46			
Gender	0.02	(0.47)	0.97			
R^2		0.55			0.55	

difference reinforces the importance of including potential covariates such as $Course$ in the analysis. The pre scores on FCI significantly predicted post scores. No other factors were statistically significant in our model.

9.3 Discussion

In our introductory physics labs, we found that students enrolled in labs that used a conceptual inquiry-based curriculum demonstrated the same decrease in the level of expert-like thinking on an assessment of their attitudes toward experimental physics as students who were enrolled in labs that used a traditional, highly-structured curriculum. However, students who were enrolled in a lab that used the conceptual inquiry-based lab curriculum supplemented by additional reflection question avoided such a decrease. A similar pattern of non-decreasing E-CLASS scores has been reported by other researchers for “open-ended” labs, labs that “develop lab skills”, and integrated approaches to lab curricula such as ISLE, SCALE-UP, or Studio Physics [309, 311, 312]. These results suggest that inquiry-based labs that include reflection on the nature of science and ways of knowing in experimental physics may also be suitable for colleges and universities that seek to transform lab instruction while attending to students’ beliefs and attitudes related to experimental physics.

In designing the additional reflection questions for our conceptual inquiry-based labs, we sought to align our students’ lab-work with the goals we identified for the course. Since we decided that the AAPT lab recommendations [1] were well-aligned to our lab goals, we looked through the recommendations and wrote reflection questions that asked students to think about topics from the lab recommendations that were not otherwise addressed in our lab curriculum. The first six reflection questions aligned neatly with E-CLASS items. Students who enrolled in the lab with the RealTime Physics + Reflections curriculum scored 1.47 points better on these items than their peers who enrolled in the lab with the Traditional curriculum. However, since the overall E-CLASS benefit from the RealTime Physics + Reflections curriculum was 3.80 points, compared with the Traditional curriculum, the results suggest that the effect of the novel curriculum was somewhat more than students simply responding to narrowly framed prompts. The process of revising lab-work to align with course goals may worthwhile for other instructors who seek to improve or transform their lab courses.

We also found a statistically significant difference between the post E-CLASS scores of men and women, when controlling for pre E-CLASS scores. This result aligns with previous findings from a large-scale study of E-CLASS scores [314], which found that women in first year physics labs who were not physics majors had lower post E-CLASS scores than men, controlling for pre E-CLASS score. Qualitative research on this topic suggests that gendered interactions between students within the masculinized culture of physics may be responsible for students having different learning experiences depending on their gender [61, 62, 74, 77, 81, 115, 141, 201, 236, 237].

Past investigations of the impact of inquiry-based lab instruction using the FCI have either reported no significant difference between traditional and inquiry-based instruction [226, 243] or studied labs in which students were simultaneously enrolled in a physics lecture course, making it difficult to determine the impact of the lab by itself [278, 294]. Our results suggest that students enrolled in a RealTime Physics conceptual inquiry-based lab had post FCI scores that were 2.60 points higher than students enrolled in traditional labs. This increase is comparable to (and is in addition to) the 5-6 point increase that students typically achieve in our introductory physics lecture course, considering that only the first half of the lab course deals with mechanics.

We note that the E-CLASS scores of students enrolled in a traditional lab class decrease during the course of the semester. For the E-CLASS, it may be possible to attribute this retrenchment to students coming to see that lab-work is less authentic than they had originally imagined, or less authentic than other labs (e.g., biology or chemistry) they have taken in college. For example, students may come to agree with the item, "The primary purpose of doing a physics experiment is to confirm previously known results," [321] from E-CLASS, shifting to novice-like thinking because their experimental work primarily seems focused on confirming theory they have previously learned in class [144].

There are several important limitations to this analysis. The E-CLASS provides only a superficial measure of students' epistemological views about experimental physics. It is not clear the extent to which students' views are being deeply or enduringly shifted as a result of the reflection questions and/or the use of a conceptual inquiry-based curriculum. These data also represent only one selective institution, with a particular set of policies for running the lab course, which may limit the extent to which these results may be extrapolated to other institutions and introductory lab configurations. Two approaches to lab-work that were not explored are the Labatorial [3, 159, 275], which synergizes lab and recitation in a way that might produce similar results to lab-lecture integrations, and online lab-work [216].

Our results suggested that students enrolled in a lab that used a conceptual inquiry-based curriculum may have demonstrated the same decrease in expert-like attitudes toward experimental physics as students enrolled in a lab that uses a traditional physics curriculum. However, when the conceptual inquiry-based curriculum was supplemented with reflection questions that addressed the nature of science and ways of knowing in experimental physics, students' attitudes toward experimental physics remained stable. In addition, our results suggested that students enrolled in a stand-alone lab that used a conceptual inquiry-based curriculum may have demonstrated improved understanding of concepts in mechanics.

Instructors who seek to transform their introductory physics labs have a wealth of curricular approaches available. Integrated lecture and lab approaches such as ISLE, SCALE-UP, or Studio Physics have been shown to improve student understanding of physics concepts as well as avoid a decrease in their attitudes toward experimental physics. "Open-ended" [311] labs and labs that "develop lab skills" [312] are likewise effective at keeping students' attitudes toward experimental physics stable. Our results suggest that conceptual inquiry-based lab curricula such as RealTime Physics may have the potential be similarly effective, if some small modifications are made to ensure that students reflect on issues relevant to experimental physics.

10.0 Expansive Framing for Curriculum and Lesson Design

What does it mean to learn something? One answer involves the concept of transfer, or the ability to take knowledge learned in one context and apply it in a different context [257]. Based on the goal of improving transfer, expansive framing is an approach to curriculum and learning activity design that brings into focus the need to situate learning and learning contexts within the broader scope of learners' settings, roles, disciplines, and experiences [94].

To understand expansive framing, it may be useful to start off with its antithesis, bounded framing. Learning activities that employ bounded framing presume that the concepts students learn are relevant only for limited contexts. These limited contexts might include specific places, times, and participants. For example, physics learners might perceive that Newton's laws of motion apply to physics problems, but are not relevant in the physical world, in their other classes, or in their future studies or career. Such a perception might develop if student learning concentrates on solving problems set in artificial contexts, regardless of the instructor's intentions or their own perspective that the laws of physics are general and broadly applicable. Bounded framing may also limit the intellectual role played by students, situating them at the periphery of the learning process [87, 242].

By contrast, expansive framing promotes student understanding of concepts by connecting between different contexts, developing links between settings and roles as a way to create intercontextuality [95]. Intercontextuality is believed to empower learners to make connections and transfer knowledge between different learning contexts (including time, location, and participants), roles, and topics. Intercontextuality supports transfer by helping learners make connections from the learning context to the transfer context by way of the encompassing context. In this view, if student learning is supported with expansive framing then students may begin to make connections between the content, the learning context, and the encompassing context. Later on, when students are asked to transfer their understandings, the intercontextuality makes it easier for them to connect ideas from the learning context and encompassing context with the transfer context [94].

In an experiment with high school biology students, Engle, Nguyen, and Mendelson found that students who were tutored with an expansive framing demonstrated substantially better transfer of their learning to a new context [95]. In this experiment, students were provided with tutoring about the cardiovascular system on one day, and then asked to transfer their understanding to the respiratory system on another day. Students received the same tutoring, but different kinds of framing. Students in the control group experienced tutoring that was framed in a typical way, while students in the experimental condition experienced tutoring with expansive framing that focused on context, topic, and roles. When interacting with students in the experimental condition, the tutors provided an expansive framing to the context of the tutoring by describing the experiment as a multi-day study (rather than two separate days), located at the university (rather than contained in the specific room), and conducted with a team (rather than with just one tutor). The tutors also described the topic of the study as "body systems" (rather than the cardiovascular and respiratory systems separately) and they emphasized that the participants were authors responsible for their own ideas (rather than recipients of ideas from others). These modest changes to the framing of the learning scenario produced dramatically improved transfer from students [95]. Another study by Engle found that an expansive framing helped 5th grade students on a science lesson [93]. In this case, the two important aspects of expansive framing were temporal connections with other contexts and the roles of the learners as members of a larger community of people interested in the topic.

Related studies in physics education have analyzed the roles of framing and scaffolding [56] when students solve isomorphic problems [186, 187, 268, 269], categorize problem types [47, 207, 208, 270], and self-diagnose their answers to quizzes [209, 317, 318]. The results of these studies serve to underline the difficulty of knowledge transfer for introductory physics students while suggesting that both scaffolding and framing could play a valuable role in improving transfer [185, 188, 198, 266].

Expansive framing can be a useful approach to interdisciplinary education. By bringing a focus to transfer and intercontextuality, expansive approaches to lesson and curriculum design encourage educators to think about how learners can make meaningful connections between the physics they learn in their classes and their personal interests and career goals [241]. Likewise, the nature of intercontextuality calls educators to ask increasingly fundamental questions about the learning goals of their courses, which may result in questions and issues that are broader than the scope of any one course, or even any one discipline.

In this chapter, we will outline how expansive framing was used in the design of student learning activities in an introductory physics lab course. First, we will consider how expansive framing was used to guide the development of lab-work. By seeking to make learning meaningful for students, many of whom were on health-science career tracks, we created opportunities for students to demonstrate transfer. This included both bringing ideas and skills from their other studies and interests into the physics lab as well as applying physics concepts in the context of their other studies and interests. Second, we will examine how we used expansive framing to improve student learning about the nature of science during the lab course. As a fundamentally interdisciplinary topic, the nature of science is a good example of intercontextuality. This allowed for ample opportunity for students to engage in reflecting on elements of the nature of science in the context of the physics lab as well as in other contexts.

10.1 Physics Labs

The introductory physics lab has long been a cornerstone in college education [64]. Traditional approaches to physics lab instruction rely on highly-structured experimental work [31, 50]. In this approach, students carefully follow instructions in a lab manual to conduct an experiment that has been designed for them. One aim of these highly-structured labs is to give students practice collecting and analyzing data in order to verify theoretical predictions. However, many students fail to see the larger goal, focusing instead on completing the assigned work step-by-step as quickly as possible, or as diligently as possible, in order to obtain their desired grades and leave. These students have adopted a bounded framing. They view their experimental work as something that is done in the lab, for a limited time, and for a limited purpose. They see their role in the lab as a procedure-follower rather than a knowledge-creator, and they are unlikely to make connections between the work they do in the lab and the physics concepts they study in class, other lab-work, or other disciplines of study.

One alternative approach to highly-structured introductory labs is skills-based labs. These labs omit the reinforcement of physics concepts as a goal, and instead focus on helping students develop their scientific thinking skills through experimental work [21, 96, 140]. The skills-based approach may allow for expansive framing, for example, if students are reminded that scientific thinking skills are useful in other scientific disciplines.

Another alternative to highly-structured introductory labs is conceptual labs. In conceptual labs, experimental work is deployed to help students learn and practice their understanding of physics concepts [278]. For example, students may develop hypotheses based on physics principles and then test their hypotheses immediately using simple, hands-on equipment. Concept-based labs may also be conducive to expansive framing, as they may allow connections to be made between lecture and the lab, between different labs, and even between the world of conceptual physics and the external world.

Our introductory lab is a one-semester, 2-credit course that is offered separately from the two-course sequence of introductory physics lectures. Students work in pairs (occasionally triplets, if there is an odd number), with up to 24 students per lab section. Each lab section is instructed by a graduate student teaching assistant (TA). Since health science majors are required to take the physics lab, but engineering students are not, the majority of the students who take the lab are interested in pursuing careers in the health sector. A smaller number are physical science majors. For both health science majors and physical science majors, the introductory physics lab may be the only (or the last) time these learners will encounter experimental physics. A few physics majors also take the lab, but they will take further lab courses in physics. For these students, the value of the lab is in how physics concepts and ways of thinking relate to the students' own studies, interests, and lives. Thus, we seek to include interdisciplinary learning in the lab learning activities and curriculum through the use of expansive framing.

10.2 Biomedical Applications

Since many of the students in our introductory labs are pursuing health science careers, the traditional, highly structured, labs were designed to include biomedical applications where possible. Typically, students would conduct experiments using traditional physics apparatus and then, for the last section, apply the same physics concepts to a biological context. For example, after a lab on physical optics and lenses, students used a model of the human eye with a water-filled "lens" that could be expanded or contracted to see how accommodation works in human vision. Another example is using a blood pressure cuff to measure blood pressure when a human arm is held at different heights, as part of a lab about Bernoulli's Principle. Students also took measurements using an EKG as part of a lab about DC circuits, briefly examined the relationship between the length of their legs and their natural gait at the end of a lab about simple harmonic motion and pendula, and learned about the theory of colors by examining spectra using a diffraction spectrometer. These activities are described in [50]. Activities such as these have been shown to foster student interest [107].

These elements of the lab work contain implicit expansive framing. The message they convey is that physics is applicable, and undergirds, phenomena from biomedical contexts, too. Students are encouraged to connect expertise from other disciplines with their studies in physics. This helps students appreciate the intercontextuality of the physics concepts from these labs, including the idea that lenses, pressure, electric currents, and simple harmonic motion are concepts that manifest in a variety of contexts.

In semi-structured interviews conducted at the end of the semester, we asked students to discuss the labs they remembered best or appreciated the most. These biomedical applications came up frequently in their responses. Ray, a neuroscience major planning to apply to medical school, described the model of the human eye as his favorite lab, saying, "we used the set-up of the eye to see where the light would focus for hyper and myopia. That was pretty cool." Zara, an anthropology major planning to apply to medical school, explained that she liked the eye model best because of the connection to a context of interest to her, explaining that "learning about the eye kind of overlaps with the anatomy that I'm interested in." Ray and Zara vividly remembered this lab experience because it connected to their interests, allowing them to develop an intercontextuality that included the concept of a lens in both physical and medical contexts.

Liza, another pre-med student, preferred the blood pressure lab activity. "I really liked when there was an application to healthcare-related things. So, like, the lab when we did the blood pressure, that was really cool." Mira, a biology major from a family of doctors, also liked the blood pressure lab, explaining that she had "never thought of how raising or lowering your arm could affect the blood pressure." Meanwhile, economics major and pre-medical student Kamala described liking the EKG lab activity, saying "I was really excited. PQRS waves. I know what valve was closing to, I knew the physiology behind it, so seeing an EKG was really exciting." Liza, Mira, and Kamala valued the biological applications because these applications allowed them to view their role as a learner expansively. They were able to connect contexts and leverage their interest and expertise in medical topics in order to gain deeper understanding of the activity they were doing in the physics lab.

However, just because a lab activity had an interdisciplinary connection to biology or medicine did not mean that the activity empowered students to frame their learning expansively. Expansive framing requires more than just an activity that is related to biomedical applications. For example, none of the students we interviewed remembered the activity in which they measured the length of their leg and period of their gait. In that activity, which aimed to help students see their legs as physical pendula while they walked, students were not provided any clues or connections as to how the activity was supposed to connect to their interests in medicine, the kinesiology of the human body, or their experiences in other science classes. Instead, it was perceived simply as a strange activity with no apparent relevance to their lives or interests.

On a similar note, Kamala, who liked the EKG activity, reported being frustrated by an activity involving color theory that came at the end of a lab in which students used a diffraction spectrometer. She saw no connection between the manipulation of light that she was doing on her lab bench and color formation by mixing light. Nor did the activity help her view the spectrometer lab expansively. Thus, it is important that lab activities seeking to provide students with opportunities to engage with physics concepts expansively be presented in a way that allows the students to make connections to pre-existing knowledge from another context, such as another class or an outside interest or expertise, or with another role, such as the students' self-concept as a pre-med student. Simply including ostensibly-relevant activities in the lab curriculum is not enough. Instead, as we saw in the case of the eye model, EKG, and blood pressure activities, the benefits of expansive framing can only be realized when learning activities include clear connections to the encompassing context, such as students' interests.

Finally, it is worth noting that expansive framing in introductory lab courses is not limited to connections with students' career plans. Janet, another pre-med student, described not a biomedical lab activity, but a simple practical lab activity when we asked her to describe her favorite lab experience. "I really enjoyed the first lab, the roller coaster. You have a thing, you drop a metal ball. Those are really fun because you see the effect of gravity on different weights of the ball. I think something that could be better is, instead of just testing two balls, what if we tested different materials? Like, how does the material affect the weight and all that stuff, and what could you predict as the distances, and stuff like that. That's something that's interesting." Her explanation of why she liked the standard ball-on-a-ramp lab demonstrates a kind of expansive framing that focuses on her curiosity and, ultimately, her role as an inquirer in the lab context. Expansive framing that situates students in the role of an investigator is eminently suited to the physics lab, including both the skills-based and conceptual labs described in the preceding section.

10.3 Nature of Science

We recently ascertained a need to do a better job of addressing the epistemology of experimental physics, sometimes called the nature of science [182], in our introductory labs. In part, this is a response to the AAPT Recommendations for the Undergraduate Physics Laboratory Curriculum [1]. We were also inspired by recent literature calling for introductory physics labs to help students learn critical and scientific thinking skills [21, 308]. The new nature of science dimension in our lab curriculum calls for students to learn about the significance and functioning of hypotheses, experimental design, correlational thinking, uncertainty, error, and the social dimension of scientific knowledge generation.

The nature of science is necessarily and unavoidably an interdisciplinary topic. Since it would be difficult to coordinate instruction about nature of science across disparate college departments, we are left to our own devices to figure out how to discuss this interdisciplinary topic in the physics lab. An expansive framing could provide the answer. An expansive framing elevates and recognizes students' expertise from other disciplines alongside their extracurricular interests. To help students learn the nature of science in the physics lab, then, we need to help students make explicit connections with their other courses and interests when we address nature of science questions. These connections, taken together, help to form an intercontextuality that both brings life to interdisciplinary essence of the nature of science while also helping students learn the topic by leveraging their expertise and interests from beyond the physics lab.

In our first attempt to introduce these concepts to the lab, we asked the graduate student teaching assistants (TAs) who taught the lab to take the lead. They were to reflect on their own understanding of the nature of science and then initiate conversations with students in the lab during 1-on-1 conversations each week as the topics arose organically through the lab work the students were performing. While the TAs initially showed interest, our observations and follow-up conversations indicated that the TAs did not initiate conversations about the nature of science in the lab, despite having opportunities to do so. They found it awkward to start such conversations when the students were focused on completing their lab work.

We may understand why this attempt was unsuccessful by considering the lack of an expansive framing lens in the design of this attempt. Because of its interdisciplinary essence, the nature of science inhabits a broader context than the lab work that students perform. We asked the TAs to invite students to engage in discussions in a context (i.e., experimental science, broadly) that was significantly different from the narrow context of their lab work, but without providing any framing to encourage or support such discussions. In retrospect, we could see that the lab work the students performed was being understood with a bounded framing. The students thought of the experimental investigations as requiring skills and understandings that were only relevant within the narrow context of the physics lab, and adopted the limited role of experiment performer, following the instructions without pausing to think about larger issues.

Our second attempt was more successful. We re-wrote the lab curriculum to include explicit questions about the nature of science that students would answer during their lab work. One example question asked students to reflect on the role of predictions: "So far, in this lab report, you have made several different predictions. These predictions have been in the form of written text, graphs, and diagrams. Why are predictions so useful and important in experimental sciences like physics? Please write a paragraph response." We included one or two such questions each week, covering the breadth of the nature of science topics that were important for us, and including space for students to draw on their expertise from other disciplines and interests. In reviewing student lab reports, it is clear that students engaged deeply with these questions. Typical responses were reflective, thorough, and demonstrated a deep understanding of the topic.

Embedding nature of science questions in the lab work, along with explicit reinforcement from the TAs that these questions were valuable, served to expand the framing for the lab work to include the nature of science as an included topic for the physics lab. As the semester progressed, students began to demonstrate transfer from other contexts to the physics lab by using examples from other science labs, courses, and disciplines in their answers to the nature of science questions. For one nature of science question that asked students to reflect on the role of the terms "theory" and "law", and on the significance of named laws (like "Snell's Law") in science, the majority of student responses included comparisons to the use of the terms "theory" and "law" in other scientific disciplines and the structural issue of the scientific establishment failing to adequately recognize the work of scientists from historically-marginalized groups, like Rosalind Franklin.

In parallel with the embedded nature of science questions, we also introduced a system of checkpoints in the lab work. When students reached a checkpoint, they called the TA over. The TA asked the students one or two open-ended questions, in order to prompt some broader thinking about scientific skills and concepts (including the nature of science). From an expansive framing point of view, these checkpoints served to reinforce both expansive contexts by asking questions that called students to consider the role of physics beyond the scope of the lab investigation they were conducting at that moment, in the physics lab. The checkpoints also served to provide an expansive framing to the role of the students in the lab. Instead of merely following the directions in the lab manual, the students were expected to take on the role of a thoughtful, reflective, knowledge generator.

10.4 Conclusion

As a tool for developing learning activities and introducing curricular transformations when interdisciplinary learning is a key priority, expansive framing may help student learners develop understandings that they are able to transfer to new contexts. Biomedical-related activities in our labs helped students frame physics concepts like blood pressure, lenses, and electric circuits in an expansive way. The framing empowered these students to transfer their knowledge both into and out of the physics lab. Likewise, when we used expansive framing to guide the introduction of explicit reflection about the nature of science into the lab course, students developed intercontextuality and were able to make thoughtful connections to other scientific disciplines. Thus, through expansive framing, students in our labs were able to engage with effective and meaningful opportunities to engage in interdisciplinary learning.

11.0 Professional Development for Physics Lab TAs

At large research universities where introductory physics labs are often taught by graduate student teaching assistants (TAs), these TAs may not receive adequate professional development for their work [63, 110, 117, 162, 184, 199, 200, 203, 270, 315]. We are seeking to understand and develop a model for how TAs can most effectively help support student learning in physics labs, and what type of professional development can most effectively help TAs facilitate lab activities so that all students can learn. Our goal is to establish a lab TA professional development program that teaches TAs how students learn in an experimental physics setting, empowers lab TAs to improve their instructional practice, and ultimately produces long-overdue learning and attitudinal outcomes [143, 169, 237, 306, 313, 323] for all students in an inclusive and equitable learning environment.

Past work has identified several important elements of effective TA professional development [252], including adapting good ideas to one's local context [154], establishing a purposeful community of practice [139], and focusing on the development of the TA's beliefs and identity as an educator [114, 121]. Another key issue is 'buy-in': TAs who do not believe in the value of the learning activity will tend to implement it with low fidelity [316], which generally negatively impacts student learning [170]. Achieving buy-in is a complex effort that depends on the context of the professional development and a variety of social cues that, when effective, work together to help TAs come to value the planned learning activities [112].

We adopt the cognitive apprenticeship model [56] as a theoretical framework for understanding both the evolution of TA learning in our professional development program and the nature of student learning in the labs. In this view, we recognize that learning requires three stages: modeling, scaffolding and coaching, and weaning. First, TAs in our training program need to witness explicit modeling of the desired outcomes. Second, learning requires careful scaffolding that supports evidence-based active engagement, and so TAs should be coached and provided guidance and support in learning how to provide this type of assistance. And third, this scaffolding and support should be gradually removed, giving TAs opportunities to practice independently. Thus, we understand TA professional development to be effective if our TAs learn about and employ effective strategies for supporting student learning, and if our students demonstrate elevated educational outcomes as a result of their TA's support.

11.1 TA Professional Development

At our large research university, approximately 33 introductory physics lab sessions are run every year. In each 'cookbook'-style lab, up to 24 students work at lab benches, typically in pairs. Labs meet for 3 hours, once per week, for one semester. Although assignments vary, typically 5 to 12 TAs are assigned to teach the introductory labs in any given semester. For many, it is their first time teaching a lab course.

Some professional development is already provided for graduate student TAs in the physics department at our university. However, all of this professional development assumes that the TAs will be small-class lecturers or recitation leaders, rather than lab instructors. This professional development includes a day-long workshop that focuses on TAs' formal responsibilities as employees, a 3-hour workshop designed to give them strategies to effectively lead recitations, and a one-credit course that teaches about effective pedagogy and affords practice with recitation-style work.

Thus, while graduate student TAs receive a variety of instructional professional development at our university, there was nothing designed specifically for lab TAs. This was a concern because some of the skills, attitudes, and approaches that are required for TAs to be effective in lab settings are not the same as those needed in recitations. Additionally, after their first semester in graduate school, most TAs are unlikely to receive any formal professional development. To rectify these deficiencies, we designed a new professional development program for lab TAs that started in the fall 2018 semester and was replicated in the spring 2019 and summer 2019 semesters. These sessions involved students using 'cookbook'-style labs that are, in general, not very effective in promoting student learning [278]. Nonetheless, while we transition to an inquiry-based curriculum, we wanted to investigate the impact of professional development on lab effectiveness given this constrained setting.

In our program [71], lab TAs meet weekly on Friday afternoons, including the Friday before the first week of classes, to prepare for the forthcoming week's lab, and to learn and practice lab-relevant pedagogy. The meetings wrap up mid-semester in order to reinforce the idea that the professional development program is a scaffold from which the TAs can be weaned, in line with our cognitive apprenticeship model, and that we expect TAs to continue developing as educators beyond our professional development program. The program was developed via extensive discussions and iterations between the author and his collaborators, all experienced physics educators.

We identified early on that motivation would be key to making this professional development program effective. Following the interest framework of Hidi and Renninger [132], we developed learning activities for TAs that would trigger situational interest in their work as TAs, reinforce that interest through meaningful social reflections, and consequently establish sustained individual interest in becoming effective instructors in their labs and beyond.

Situational interest reinforcement happens at the start of our weekly meetings, when all the TAs share a student interaction from the past week they found surprising, concerning, or encouraging. These reflections provoke cross-discussion in which TAs celebrate their progress as educators, reaffirm shared commitments to helping students learn meaningfully, or brainstorm approaches to uncommon problems. These discussions are moderated by the training leader, using standard methods for establishing and maintaining positive interactivity [53]. At all times, we focus on giving TAs opportunities to speak out in order to develop their confidence as educators.

Most of the lab TA meeting is dedicated to one or two relevant learning activities, which are designed to promote TAs' individual interest as educators. The learning activities are intended to help TAs directly improve their skills at working with students, better understand the nature of student learning [233], and develop both proficiency with the apparatus and increased levels of motivation to support students.

One learning activity we employed is role-playing student-TA interactions around key points in the lab. For this, the trainer sets up 'sabotaged' experiments in which the apparatus is miscalibrated, the analysis is incorrectly done, or a similar common issue. The scaffolding in this experiment allows TAs to practice interactions that support evidence-based active learning. After role-playing through an interaction, the TAs and trainer debrief and move to another experiment. Two other approaches we used occasionally are demonstrations, which serve as models for TAs to replicate, and conducting carefully-moderated whole-group discussions. Some sample learning activities are listed in Table 17.

Table 17: Sample of learning activities in the lab TA professional development program.

Activity	Type	Weeks
Icebreaker	demonstration	first
Reflections	group discussion	all
'Sabotaged' experiments	role-playing	most
Equitable learning environments	group discussion	third
Teaching about the nature of science	group discussion	fifth

11.2 Methodology and Results

We assess the impact of our lab TA professional development program by analyzing three different examples of ways in which the training program might have impacted TAs or students. For clarity, the methods and results of each of these three examples are presented together. We begin by asking whether the program changes how TAs think about their work in supporting student learning. Next, we investigate the behavior of TAs as they work in their lab sections to understand if the professional development has changed the nature of their interactions with students. Last, we ask whether the program has a 'second-order' effect by improving students' learning outcomes.

11.2.1 TAs' Attitudes Toward Student Learning

To assess the direct effect of the professional development program, lab TAs were asked to write short responses to variations on the question, "How will you help students have effective learning experiences?" at the start and end of the program. In total, 13 responses were collected at the start, and 11 at the end, of the fall and spring programs.

The responses show a marked transformation in how the TAs viewed their role in helping students to learn. At the beginning of the professional development program, most of the TAs described their role in terms of a traditional transmission model of education [69], as these representative excerpts illustrate:

> I will explain to them which equipment corresponds to which concept. Then they can build connections of the physics concepts to the lab.

> [I will] add in some physical explanation into the demo at the start of lab... a feeling on physics will build up subconsciously after they leave the lab.

The use of the verb 'explain' is abundant in these early responses, as the TAs view themselves as either telling students about physics concepts or clarifying the procedure for the lab-work. Other responses emphasize the importance of good lecture structure, clear expectations for lab report formatting, and creating a "relaxing" environment for the students. Given that most of the TAs have experienced a traditional, transmission-based style of education – and are continuing to experience that model as graduate students – it is not surprising that they rely on the transmission model of learning to frame their work. Likewise, it makes sense that the TAs view their role as helping to make the lab easier for students, as that is likely how they viewed their own TAs during their recent undergraduate experiences.

By the end of the the professional development program, however, the emphasis shifts and most of the TAs' responses celebrate students figuring things out on their own and with their lab partners, as seen in these typical quotations:

> I would encourage a student... to collaborate with peers, ask themselves more rigorous questions, etc.

> I asked her to think of the problem in a physical sense, instead of plugging in given equations. She actually came up with the right answer... by herself.

These final reflections indicate that after the professional development the TAs have generally discarded the transmission model in favor of a student-centered view of learning, which was one of our goals. 'Explain' is no longer used, and the responses tend to center the student's experience instead of the TA's work. Other responses emphasize the importance of encouraging positive collaboration between students and explain techniques the TAs use for supporting student meaning-making without directly supplying information.

It is not possible to conclusively determine what caused the shift in TAs' views about learning. Was it the professional development program? Their experiences as a TA in the lab? Something else? However, by comparing experienced TAs who have taught the lab before (in semesters before the professional development program) with newer TAs who have never taught labs before, it seems likely that the common factor – the lab TA professional development – played at least some role in their movement from a transmission to a constructivist view of learning.

11.2.2 TA-Student Interactions

Since one goal of the lab TA professional development is to get TAs to help students frame and answer their own questions, rather than just offering advice and explanations, we hypothesized that there would be a change in the nature of student-TA interactions after introducing the program. To measure the extent of these changes, we used the Real-time Instructor Observing Tool (RIOT) [303] to categorize these interactions for the same group of TAs.

The RIOT allows an observer to continuously categorize the types of activities undertaken by an instructor, in this case the lab TA. When the instructor switches from one type of activity to another, the observer records the nature of the new activity along with a timestamp. Other than infrequent occasions when the TA would be briefly checking personal notes or be out of the room, the seven categories shown in Fig. 9 capture the full breadth of TA activities during the lab sessions that we observed. The data presented in Fig. 9 depicts interactions for 13 TAs over 99 hours of instruction. We chose to observe during weeks 3 and 6 of our 13-week semester because the lab interactions should have reached a 'steady

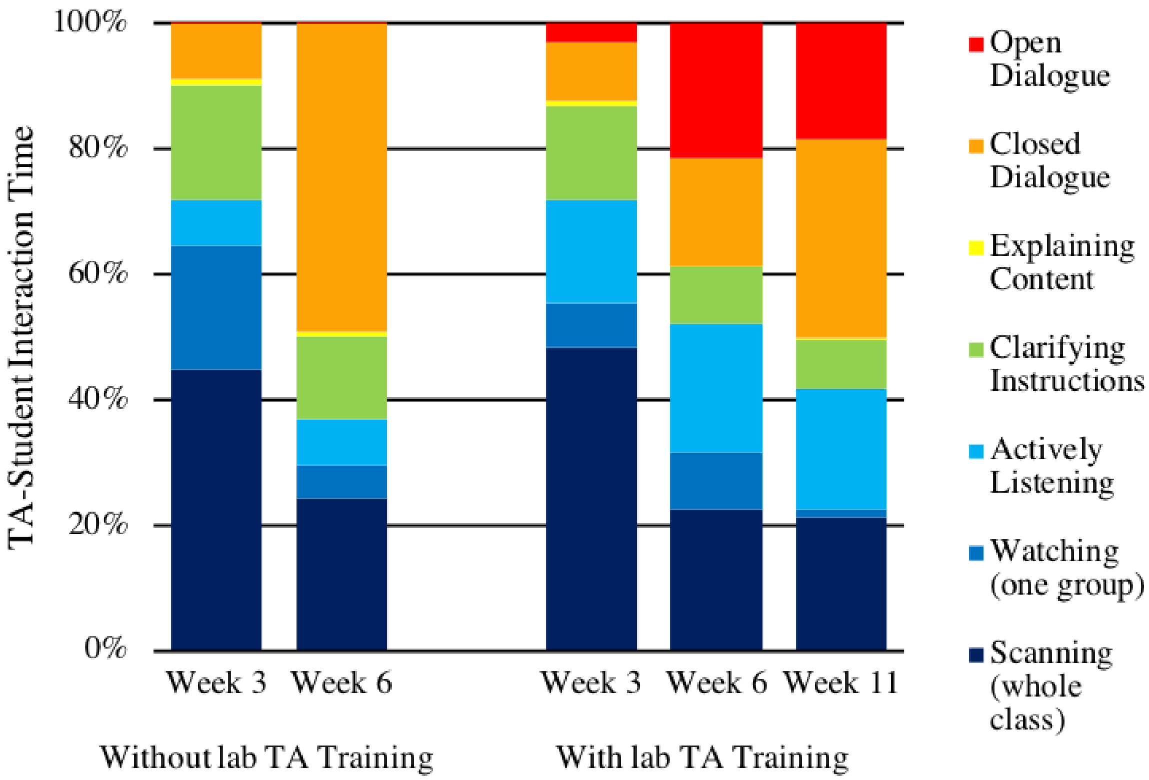

Figure 9: RIOT data from a semester before the introduction of a lab TA professional development program compared with data from the fall 2018 semester, when the program was implemented.

state' and because the lab-work for those weeks was typical, not excessively intricate, and didn't require that the lights be turned off. We also observed during week 11 for the TAs that received the professional development to explore if the effect of the training diminished after the meetings wrapped up in week 6. Complete definitions and examples of the seven types of interactions we observed are available in Ref. [303].

We adopted slightly different labels for some categories for clarity, but use the same definitions and meanings for these interaction types. For example, we use 'Actively Listening' (to question) to emphasize that the TA is engaging with non-verbal listening cues, unlike in the case of 'Watching'. Likewise, 'Explaining Content' and 'Clarifying Instructions' are forms of "talking at students" [303]. We exclude several interaction types that are not relevant to our labs and the relatively rare cases when the TA is not interacting with students.

Two specific examples of these categories will be relevant to our analysis. 'Open Dialogue' is an interaction in which a student is contributing more than half the words (or ideas) to a conversation and is actively developing an understanding of physics or lab ideas, while the TA plays a supporting role by asking prompting questions or helping students to frame their ideas, as opposed to 'Closed Dialogue' which is TA-dominated. 'Actively Listening' means that the TA is near a group of students and showing an active interest in their discussion through non-verbal cues such as establishing eye contact, body positioning, or gestures, but isn't participating in the conversation as a contributor.

Two differences stand out in Fig. 9. First, the amount of time that TAs devote to 'Open Dialogue' is larger for the TAs who have completed the professional development program, and it seems to increase while the TAs are engaged in the program. This behavior aligns with our goals and with the responses we discussed in section IIIA: the TAs are more likely to let students lead their conversations, prompting rather than telling, and helping the students to formulate and answer their own questions.

Second, the TAs who have taken the professional development program devote approximately twice as much time to 'Active Listening' as those who haven't taken the program. In practice, this manifests as the TAs being more willing to engage with students and offer non-verbal support and attention as the students complete their lab-work. In our observations, this increased level of 'Active Listening' appears to be aligned with TAs being increasingly comfortable interacting with students when they can position themselves as guides rather than feeling compelled to adopt the role of an all-knowing expert. In other words, the TAs were comfortable simply being there and talking with the students, and didn't feel the need to adopt a 'hide or provide' behavior, in which they only approached student groups if they felt they had some information to share.

Teaching is certainly a complex endeavor, and no one interaction type is necessarily better than any other in all circumstances. Good educators typically use a combination of many types of interaction [232]. Overall, though, we find some evidence that the balance of interaction types after the professional development program is better-aligned to the goal of supporting student-led active learning for TAs who have taken the program.

11.2.3 Student Outcomes

While both TA attitudes and interactions show improvements, the fundamental goal in our new TA professional development program is to improve student learning in the labs. Thus, in order to ascertain the impact of our lab TA program on student learning, we need to compare students whose TAs did deliver the learning support we designed in our professional develop program with those students whose TAs did not deliver this support. Here we discuss it in relation to the lab TA professional development module related to the nature of science (NoS).

The nature of science is a set of beliefs about the epistemology of science, i.e.: principles that we might identify as fundamental characteristics of Western scientific work. These include such ideas as 'theories require evidence', 'science makes predictions', and 'scientists seek to avoid bias' [7].

We observed that for the fall 2018 implementation, our professional development module for the nature of science had an abnormally low level of engagement from our TAs. During the program, most of the TAs merely went through the motions, and during lab observations later in the semester we saw no evidence that they used the proposed strategies in their labs. However, two TAs clearly and unambiguously bucked the trend: they engaged in a lengthy discussion about the value of explicit instruction on the nature of science that went beyond the training session, and we observed them using our strategies in their lab sessions on multiple occasions. Thus, by comparing the 74 students who received this NoS treatment from their TA to the 207 who did not, we estimate the impact of having a TA engaged with NoS instruction.

As a dependent variable, we rely on a categorization of E-CLASS items [308] that identified a cluster of 6 items as relevant to NoS in our context [72]. E-CLASS scores indicate the degree to which a student agrees with expert-like attitudes toward 30 items related to experimental science. As shown in Fig. 10, when we consider the 'Other TAs' that did not engage with the value of teaching NoS, the average change in score for the 6 NoS-related items is more negative than for each of the 24 other E-CLASS items, indicating that our students typically regress more on the nature of science items than on the other items in E-CLASS. And while the decrease in our setting is somewhat more negative than in the national sample from [308], a similar pattern appears, suggesting that these 6 NoS-related items are particularly vulnerable.

For the TAs who have adopted NoS instruction, we observe a decrease on the 24 non-NoS items that is comparable to the decrease seen by the other TAs, but a sizable increase on the 6 NoS-related items. Using a mixed-effect ANCOVA model controlling for overall pre-test E-CLASS score, and with satisfactory normality and homogeneity of variance, we find a statistically significant difference between our two TA groups and the two item categories ($F(1, 279) = 11.5, p < 0.001$).

We note that the lab-work in this study employed a 'cookbook'-style approach that will be replaced with an evidence-based active learning approach [278] starting in fall 2019. We expect that overall E-CLASS score improvements will require both that TAs are effectively trained (and adopt the strategies learned in this professional development) and that inquiry-based learning activities are in place.

We can draw two conclusions from this result. First, students attitudes about experimental physics, as measured by at least some of the E-CLASS items, can be influenced by the interactions and learning that are offered by graduate student TAs in the introductory lab. Second, TA buy-in for particular instructional strategies is essential in order for this to happen.

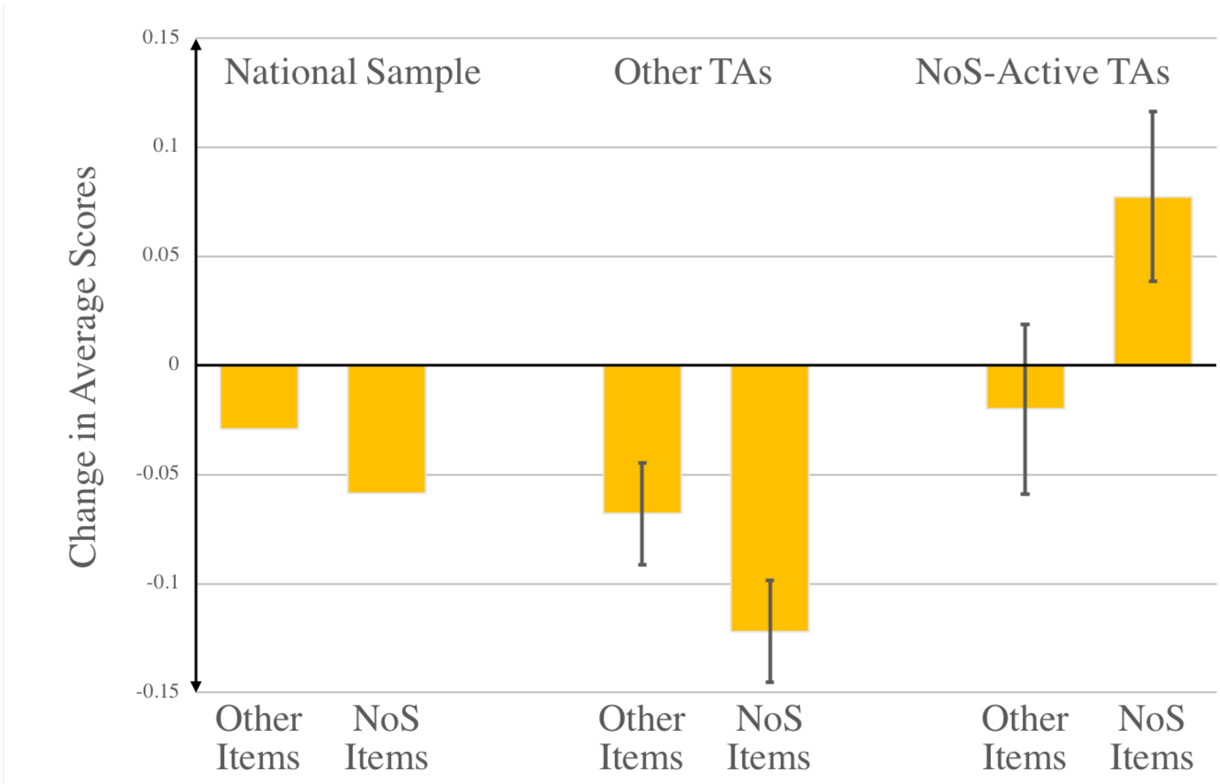

Figure 10: Change in E-CLASS item scores between pre- and post-instruction assessments, averaged over the 24 and 6 E-CLASS items identified as either not, or belonging to, a cluster related to the nature of science.

11.3 Discussion and Implications

At three levels of analysis, we find evidence that an effective lab TA professional development program has the potential to positively impact the work undertaken by TAs and the learning of their students. After the professional development program, lab TAs demonstrated a shift in how they viewed their role as instructors and how they thought about the nature of student learning. It appears that the lab TAs who completed the program were also more likely to 'walk the walk,' interacting with students in ways that are commensurate with what we sought to teach them about supporting active engagement learning. And when lab TAs took up the new approach to teaching a topic, such as explicitly engaging in discussions about the nature of science, student attitudes corresponding to that topic became more expert-like.

The nature of the work done by lab TAs, the activities involved in our program, and the improvements noted above act together to shine light on the importance of specialized professional development for TAs who are working in lab settings. While there is overlap with how we might train TAs to support evidence-based active engagement learning, the need to develop mastery at troubleshooting apparatus and to address issues such as the nature of science means that lab TAs should be receiving dedicated professional development in order to be effective in their work.

As seen in the case of our nature of science module, some improvement is still needed for our professional development program. TA 'buy-in' remains a vital issue, especially for the NoS module, and as we go forward we will continue working to improve the social and structural factors that can drive TA buy-in [112]. However, even with an imperfect professional development program and 'cookbook'-style labs, and through the inevitable noise of implementation, these results suggests that lab TA professional development has the potential to have a positive impact on TA performance and effectiveness.

12.0 Professional Development Combining Cognitive Apprenticeship and Expectancy-Value Theories

At large research universities in the USA, introductory physics labs are often taught by graduate student teaching assistants (TAs). However, prior work has shown that many graduate students may not be ready to effectively lead instruction. In a variety of settings, graduate students adopted less-effective pedagogical strategies for working with students [110, 203, 204, 202], demonstrated inadequate understanding of the nature of different problem types [117, 184, 270], and demonstrated low levels of pedagogical content knowledge when asked to identify common student difficulties [162, 199, 200].

One popular approach is to have graduate students enroll in a course that teaches them about physics pedagogy [9, 178, 196, 218, 219, 252, 282]. Prior studies have identified several key elements to effective professional development for graduate student TAs in physics [154]. Focusing on the TA's beliefs and identity as an educator is essential [113, 114, 121], as is establishing a purposeful community of practice when working with a group of TAs [139]. Two other important considerations are the importance of respecting and supporting TAs' emerging competencies as educators [9, 219] and the need to clearly align expectations for the work that TAs do with the types of tasks their students perform [295]. Maries [196] identifies three key components to TA preparation: attending to psychological factors such as anxiety about their role as a TA, providing ongoing support, and attending to their beliefs about teaching to achieve buy-in.

The issue of buy-in from the TAs is a critical one if TAs are expected to implement instructional practices with a high degree of fidelity [63, 86, 170, 315]. While graduate student TAs may express support for the purpose and goals of a particular curricular strategy, the way they employ that strategy may deviate from what was intended as they navigate the tension between the pedagogy they learn in the professional development program and the desires of their students to minimize the amount of work or "thinking" they are required to do while still achieving good grades [48, 316]. For example, TAs may seek to avoid alienating or frustrating their students, or may seek to make their students' work easier, by providing shortcuts or telling students answers [48]. In particular, while navigating these types of tensions, TAs who choose to use less-effective learning strategies that are not supported by research in physics education may cause their students to learn physics concepts and develop experimental skills less well than they otherwise might [170].

Moreover, while a well-designed and thoughtfully implemented TA professional development course may be helpful, TAs who will lead introductory lab courses may require additional, specialized preparation [9, 192], such as instruction on pedagogy relevant to the lab, at weekly lab TA meetings [282]. The goal of this chapter is to report on a particular instantiation of a lab TA professional development program developed using an iterative approach over several semesters based upon a framework that combines cognitive and motivational theories and describe how it was implemented and evaluated. The evaluation involved the use of informal and structured observations, interviews, and writing prompts for reflection. We discuss the utility of focusing on TA engagement and buy-in and demonstrate that a carefully designed lab TA professional development can have a positive impact on TA attitudes and practice.

12.1 Theoretical Framework and Research Questions

The theoretical framework for our lab TA professional development is illustrated in Fig. 11, and outlined below. This same framework was also useful in helping TAs reflect upon their students' learning in the lab, which also requires attending to both cognitive and motivational aspects.

The framework guiding our particular instantiation of the model of lab TA professional development integrates cognitive and motivational factors. We took inspiration from the cognitive apprenticeship theory [56] to structure the guidance, scaffolding, and support needed by TAs in order to flourish as educators. Employing the cognitive apprenticeship theory, our model of lab TA professional development focused on first giving TAs opportunities to observe effective pedagogical approaches via demonstration of the criteria for effective performance. Next, we provided coaching and scaffolding as the TAs practiced instructional skills

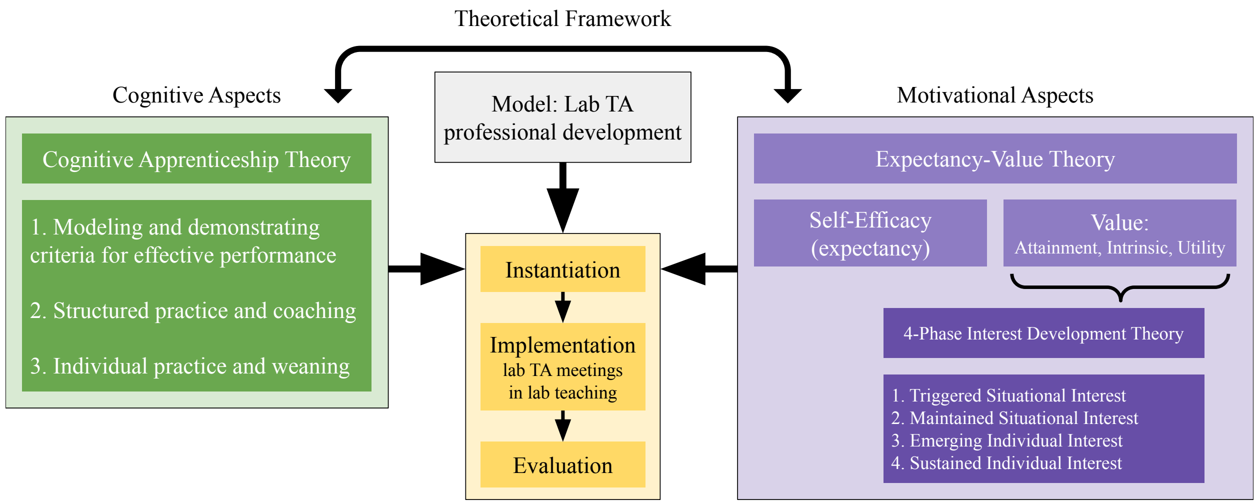

Figure 11: Framework for an investigation that uses Cognitive Apprenticeship and Expectancy-Value theories inform the instantiation of our lab TA professional development program.

and developed as effective TAs themselves. Finally, the TAs were weaned off the support to help them develop self-reliance. For example, during the lab TA meeting, TAs would observe a demonstration of a TA-student interaction and then be given time to reflect upon, and practice by role-playing, that type of interaction with fellow TAs. Individual practice would occur during the TAs' actual lab sections with students. The fact that the TAs were undergoing the professional development while also teaching the lab simultaneously was central to the success of our approach.

To account for motivational factors, we grounded our work in expectancy-value theory [92, 230]. Expectancy-value theory (EVT) posits that expectancy and value both will influence decision-making and determine, e.g., the extent to which the TAs will employ the pedagogical approaches learned in the professional development program in their actual practice as TAs. Expectancy is closely-related to Bandura's construct of self-efficacy [20], and can contribute to a graduate student's agency as a TA in implementing effective pedagogical approaches learned in the professional development program. In EVT, value includes the decision-maker's interest, attainment value and extrinsic value in pursing a given course of action, e.g., TAs engaging meaningfully in the professional development program and using pedagogical approaches learned in their actual practice as lab TAs. In particular, EVT describes ways to understand the value individuals discern in the work they do, including attainment value, whereby a TA feels that success is personally meaningful; intrinsic value, whereby the TA experiences some level of individual interest in their work; and utility value, whereby the TA is motivated by the relation of their work to current and future goals [92]. The EVT played an important role in our conceptualization of the lab TA professional development program because we were concerned with the question of whether, and how, lab TAs will choose to engage in our lab TA professional development program, to what extent they change their views about instructional practices, and how they go about interacting with their students in their lab course as a result of the program. In other words, EVT guided us to contemplate the importance of these motivational factors in our professional development program to increase TA engagement and incorporate strategies to increase TA self-efficacy and interest in their use of effective pedagogical approaches learned in their lab sections.

Moreover, we recognized that the TAs in the professional development program will initially be at different levels in terms of their interest in teaching their lab and we must structure the activities and interactions between the TAs in the professional development program as well as the interactions between students and TAs in the lab in such a way that they are productive and propel the TAs to the next level of interest and engagement. We took inspiration from the four-phase framework of interest development [132], which accounts for TAs with different levels of interest, and proposes pathways for TAs to progress to higher levels of interest. In this framework, interactions between the TAs during the professional development activities and between TAs and students during the lab sections may serve, e.g., to provoke situational interest in a TA who would otherwise not place a high value on their TA work. That triggered situational interest might be elevated to the second phase, maintained situational interest, e.g., if the TA is provided with continued support during professional development and feels confident supporting student interactions in their lab sections effectively. The third phase, emerging individual interest, may require that the

TAs begin to find personal meaning in their work as TAs. Individual interest development could be supported, e.g., by having the TAs reflect in groups and individually about how they play an important role in shaping their students' learning, gradually giving them more autonomy in coming up with effective strategies for engaging their students meaningfully in the lab and connecting their TA work with their other interests. For example, for a TA interested in equity work, pointing out their role not only in student learning but also in mitigating inequities in the lab has the potential to make them commit to their practice as a TA with greater interest. Finally, for TAs who have extended, personally meaningful practice, the fourth phase of well-developed individual interest may be attained. Illustrative examples of how this progression has been enacted by TAs in our professional development program is provided in Table 18. At a finer level, we note that TAs may not be described by a single phase of interest development theory for all the aspects of their instructional work. For example, at a given time, a TA may have an emerging individual interest in helping students make connection between their experiments and physics concepts but only situational interest in helping students make accurate measurements.

We note that as a pre-requisite for TAs engaging productively in the professional development that is offered during the weekly lab TA meetings, it is essential to get buy-in from the TAs. In particular, it is necessary that TAs believe that the professional development activities will help them to become better educators, that their own improvement as educators will help their students learn more effectively, and that both of these things are desirable and achievable. The expectancy-value theory was useful in contemplating this question of how to increase TA buy-in. We probed signatures of growth in TAs' self-efficacy and value as reflections of their buy-in via qualitative analysis, including through informal observations of their lab sections throughout the semester and individual interviews with a subset of TAs.

We evaluated our implementation of lab TA professional development in two ways. First, we investigated whether our lab TA professional development had a positive affect on TAs' views about the learning process. A TA's attitudes toward teaching and learning, and specifically whether they describe their work in terms of a transmissionist or a constructivist lens, can be a powerful predictor of learning [176]. Second, we sought to understand how our lab

Table 18: Example of how a TA's interest might evolve according to four-phase interest development theory.

Phase	Example scenario
Triggered Situational Interest	The TA sees that a student's experimental results do not follow the expected trend and wonders how to talk with the student about this.
Maintained Situational Interest	The TA continues to ponder what the students might have done to get such unexpected results. She mentions it at a TA meeting, and engages in a discussion with other TAs who noticed the same thing in their class.
Emerging Individual Interest	The TA begins to think proactively, predicting where students might make missteps while she prepares for future labs, and continues to think about how she can intervene to help students to check their own results.
Well-Developed Individual Interest	The TA takes pride in her work, enjoys interacting with students, and may begin to see herself as an educator. She seeks out challenges, such as empathetically spending additional time with a student who hasn't been asking for help but may need extra support.

TA professional development impacted the nature of TA-student interactions while the TAs guided student learning in the lab. Effective evidence-based active engagement learning, for example, that we helped TAs reflect upon and practice with other TAs in small groups calls for TAs to dedicate time to supporting student meaning-making in open-ended discussion [118, 128, 90]. TAs' attitudes toward learning and their behaviors supporting learners were evaluated, e.g., using analysis of writing excerpts and an instructional observation protocol. Our investigation focuses on addressing two research questions:

- Do TAs' views about teaching and learning change after lab TA professional development?
- Do TAs' behaviors as lab instructors change after lab TA professional development?

12.2 Lab TA Professional Development Program

Our lab TA professional development model is implemented during weekly lab TA meetings. The first meeting occurs on the Friday before classes start ('week 0'), and they continue for subsequent Fridays through the semester. The lab TA meetings are run by a senior TA (i.e., a graduate student TA who has taught the lab in a previous semester and shown potential to be successful supporting their peers), with support from the lab coordinator. Although the author served as the senior TA for the first two iterations of these meetings, we have now implemented the lab TA professional development for three semesters with a senior TA who was not involved in its development. Our aim is that descriptions and video demonstrations of the activities, combined with structural elements of the lab course such as the checkpoints, will serve to 'futureproof' our efforts when faculty teaching assignments change and TAs graduate.

The framework described above, including both cognitive and motivational aspects, serves as the theoretical underpinnings to the structure of the lab meeting as a whole, and also for the individual activities that are conducted. We attended to the following principles when designing and implementing our instantiation of the lab TA professional development:

1. When TAs learn new skills or concepts, they should be guided through the three stages of cognitive apprenticeship.
2. While TAs enter the professional development program with differing levels of self-efficacy and value related to their work in supporting student learning, all have space to grow as educators.
3. TAs will evolve in their interest (and consequently, demonstrate increased agency as instructors) according to the 4 phases outlined above if they are provided adequate opportunities to reflect on their instructional practices, learn and practice new skills, and practice using those skills in their lab sections with their students.

All the lab meetings begin with a group reflection activity in which each TA is asked to share an experience, insight, or concern from the previous week's lab. Sharing experiences may help some TAs to transition from a triggered situational interest to a maintained situational interest phase, while their peers can provide support by helping to think about solutions to problems TAs have been facing in their labs. In addition, encouraging all TAs to speak (after being given a prompt and time to come up with something to share) can be especially valuable for English language learners as it helps to normalize oral communication in the lab environment, and in the lab TA meetings specifically.

Following the reflection activity, the lab coordinator shares procedural and apparatus notes about the lab. This takes no more than 10 minutes, and is done in a way that models how the TAs can share relevant information about the procedure and apparatus with their own students. The apparatus is briefly demonstrated, and complex issues and some common student difficulties are identified.

The remainder of the meeting time is devoted to one or more activities, some examples of which are described below. Complete descriptions are available online [71]. The first lab TA meeting includes an icebreaker activity that asks TAs to briefly discuss what worries and excites them about working in the lab, as a way to promote growth in their self-efficacy (by seeing that others are worried too) and the way they value their work (by normalizing excitement about leading lab sections). An example program of activities for one semester is outlined in Table 19. The activities and program were developed, iterated, and trialed prior to their implementation, so that the evaluation below follows a relatively-mature implementation of our lab TA professional development program.

12.2.1 Sabotage Activity

In the 'sabotage' activity, two TAs engage in a role-play that helps them to better understand the apparatus and experiment while also practicing TA-student interactions in a realistic and relevant scenario. A lab station is set up with data collection and/or analysis completed in line with a certain stage of the lab procedure. However, one piece of the apparatus, procedure, or analysis has been done incorrectly, or 'sabotaged'. For example, a sensor may be uncalibrated, a dynamics track may not be level, a circuit may be constructed with a short, incorrect units may be used, data on a graph may be displayed in a confusing way, or a linear trend-line might be applied to non-linear data.

One of the participants will role-play as a student, and may need a quick briefing about this stage of the lab procedure. This is usually provided in the form of a couple of sentences on a slip of paper (e.g.: "You have collected these data indicating the position and time of a falling object, and drew a linear fit on the position-time graph using the computer. However, the line doesn't seem to fit the data perfectly, and you're not sure what the slope represents."). The other participant role-plays as a TA. The 'TA' approaches the 'student' and initiates a conversation. The TA aims to help the student resolve the issue with the experiment, but must do this using techniques that support inquiry learning (e.g., standing on the side, not touching the apparatus, only asking questions). Following the role-play, the participants reflect on their experiences, e.g., the students discuss what they found helpful, and the TAs discuss the strategies they employed to help the students.

The TAs who attend the weekly meetings appreciate this activity because it gives them practice dealing with common student difficulties, which are often the inspiration for the sabotage. The TAs also get the opportunity to practice interacting with students in a low-stakes setting, and they receive feedback from their peers, the senior TA who organizes the lab meetings, and the lab coordinator. In addition, the TAs get a chance to experience how students feel when they receive support from a TA.

12.2.2 Task Division Activity

Prior research has identified task division (e.g., based on gender) as a possible cause for inequitable work in introductory physics labs [74, 236, 237]. In this activity, TAs are shown what this type of inequitable task division looks like and there is discussion of strategies for countering it in their labs.

In pairs, the participants are asked to engage in a challenging lab exercise, such as constructing a certain circuit. After most of the pairs are finished, the senior TA shares their observations of pairs in which one participant was more engaged than their partner, and then explains that adopting gendered modes of work (e.g., women taking on secretarial roles while men do the tinkering), among others, is a common 'bad habit' [126] and that one strategy to countering task division is to re-group students once or twice per semester.

Next, the participants are randomly assigned new partners, and asked to complete another, similar task, such as building a different type of circuit. This time, they should be mindful of their own roles, and also watch other groups around the room. Once the participants finish this second task, they reflect as a group on what they saw and experienced. The senior TA shares some relevant strategies for countering inequitable task division (assigned roles, negotiating fair task splits, giving students individual opportunities to develop core skills). This type of reflection is designed to help TAs understand what inequitable task division looks like, and why it is problematic, before giving them tools to respond when they see it in their own labs.

12.2.3 Dominance Activity

Another way that inequities can manifest in group work is through domination of discussions. Based on work by Turpen, Sabo, and others [250, 293], this activity gives TAs an opportunity to observe conversation dominance, and to discuss and practice discursive techniques they can use to address imbalanced interpersonal dynamics. First, two participants (recruited in advance) act as quiet and dominant students. A third participant is chosen to act as the TA. The rest of the group observes while the TA role-plays a check-in with the two actors.

After the role-play, the two actors share their experience and the observers are invited to both comment on what they saw and suggest 'teaching moves' that the TA could make in such a situation to bring forward the voice of the quieter student. This rich, participant-driven discussion typically touches on topics such as body positioning, non-verbal cues, and tone and speaking patterns, as well as on the question of who is being recognized, praised, or ignored in TA-student interactions. TAs also discuss which student is being recognized and being given an opportunity to develop their self-efficacy as a physics person. Finally, the participants break into groups of three to replicate the scene they just observed. Each participant is given a chance to try some teaching moves as the TA, to get practice implementing them, develop their interest, and to see what works best for their own style of interacting with students.

12.2.4 Nature of Science Discussion

One aspect of the 'thinking skills' identified by our faculty is the need for students to understand the broad-scope epistemology of experimental science, sometimes called the nature of science [12]. For this activity, TAs engage in discussion to unpack the meaning of "the basic beliefs and attitudes that scientists share about what they do and how they do their work" [7], then reflect individually and as a group on how they might be able to have discussions with their students about these beliefs.

The goal here is to have TAs generate both understandings and approaches themselves, which could stimulate their emerging individual interest (or help them progress along the axis of four-phase interest development theory) while also providing meaningful, practical strategies for talking about the nature of science with their students. As reported elsewhere [73], this approach to introducing discussions about the nature of science to the lab course was generally unsuccessful, and will be a focus for future work (see below).

Table 19: Program of lab TA meeting activities during one semester of the professional development program.

Week	Activities
0	Icebreaker and norms, Writing activity,
	Overview of responsibilities (by lab coordinator)
1	Socratic dialogue activity,
	Sabotage activity
2	Nature of science discussion, Practice
3	Sabotage activity
4	Task division activity
5	Sabotage activity
6	Dominance activity
7	Task division activity
8	Sabotage activity
	Revisit strategies for supporting inquiry
9	Sabotage activity
10	Revisit nature of science
11	Sabotage activity
12	Writing activity

12.3 Methodology

12.3.1 Participants

This study reports on implementation of the lab TA professional development at our large state-related research university in the USA with a student population of 18000 undergraduate students and 12000 graduate students. Our data reflect a total of 44 sections of the introductory physics lab, led by 30 different TAs, during 2018 and 2019.

Each lab section meets for three hours, once per week, for a semester. In the lab, 24 students work in pairs at a lab bench that is equipped with a computer. In addition, and before their weekly lab section, students also attend a one-hour lecture delivered by the lab coordinator that aims to (re)introduce the key physics concepts the students will be exploring during the lab. The lab is a one-semester class that is taken separately from lecture-based Physics 1 and 2 courses, and focuses on physics concepts from Physics 1 and 2, including mechanics, electricity and magnetism, and optics, and requires completion of Physics 1 as a pre-requisite.

The lab sections were led by graduate student TAs who had completed, or were currently enrolled in, a one-semester course on physics pedagogy. 80% of the TAs were international students, 70% identified as male, and 85% were in their first or second year of graduate school. Nearly all of the data reported below are for TAs who were teaching the one-semester lab course for the first time. The high fraction of lab TAs who were international students may be the result of a policy requiring better results on an English proficiency test in order for a TA to be appointed to lead recitations. Interestingly, in our interviews with students, they rarely indicated they had difficulty understanding their TA.

Two offerings of the introductory physics lab are offered: an algebra-based version that accounts for the majority of lab sections and attracts primarily health science majors, and a calculus-based version that primarily enrolls physical sciences majors. Aside from some small differences in the presentation of the theory in the lab lecture, the lab is identical for these two offerings. For our student population, those who wish to apply to medical school (and several similar career trajectories) require a physics lab, while most engineering students are not required to take a physics lab. The algebra-based offering is about 60% female, and primarily students in their third or fourth year of study. The calculus-based offering is 30% female, and primarily students in their first or second year of study.

12.3.2 Traditional and Transformed Lab Curricula

Our observations and data follow TAs who instructed two types of introductory labs. For the first two semesters, the labs followed a traditional, highly-structured format [50]. In the traditional lab course, students read a section of their lab manual that developed the theory, then carried out experiments in a step-by-step manner that was designed to allow them to operate sometimes-complex apparatus efficiently, but left little room for students to develop experimental skills or develop their epistemological understandings of experimental science.

For the third semester of our study, the labs were transformed to an inquiry-based format. The process of transforming physics labs has been thoughtfully described in prior research [137, 322], and our work followed a similar process. We started by meeting with faculty in our department to identify goals for the introductory physics lab courses. Using the AAPT Recommendations as a guide [1], the emergent consensus was a focus on improving fundamental lab skills such as making measurements, creating graphs, and troubleshooting, and enhancing thinking skills such as understanding models, devising hypotheses, and developing scientific arguments [22].

In addition, our faculty members decided that the introductory physics lab should also aim to help students improve their understanding of essential physics concepts by having students engage with inquiry-based lab-work. For some time, physics educators have argued that the introductory physics lab should focus on inquiry, rather than highly-structured laboratory work [31, 60]. Inquiry-based lab work such as Real Time Physics [278], which we adopted as the foundation for our lab curriculum, provides scaffolding as students develop their conceptual understanding of physics concepts in the lab. We have supplemented the Real Time Physics labs with additional learning activities and questions to explicitly help students develop the lab and thinking skills our faculty members identified as priorities.

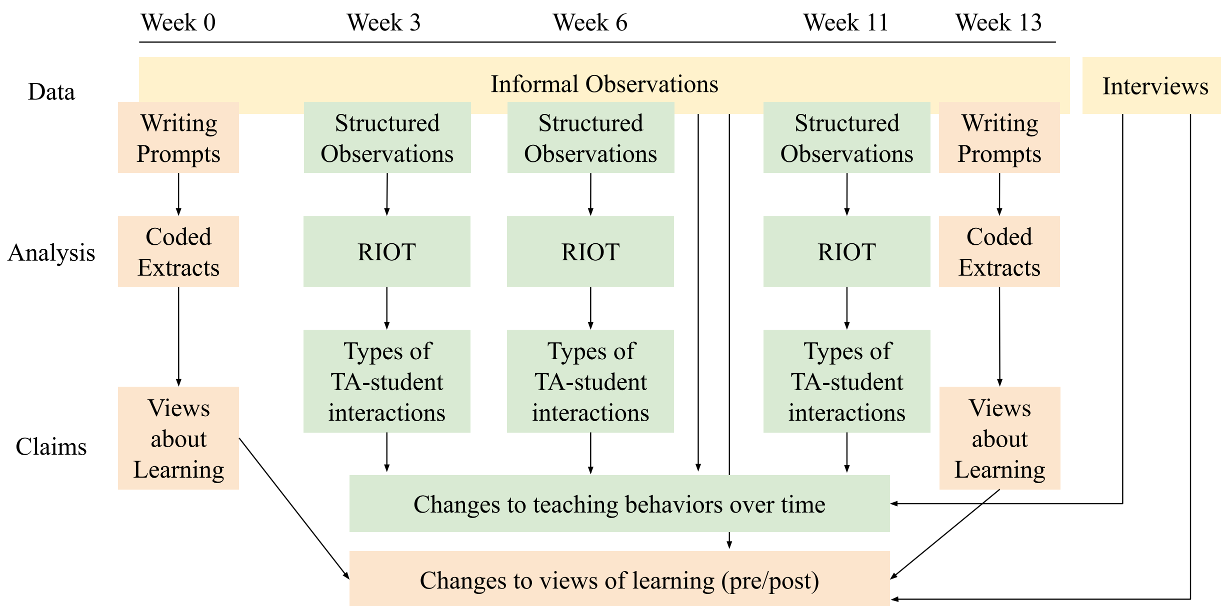

Figure 12: Diagrammatic representation of data collection over the course of a typical semester.

In an effort to improve the quality and quantity of TA-student discussions and feedback about physics concepts and lab-related thinking skills, we introduced a series of checkpoints to work that students do in the lab [153]. In particular, 2 to 4 times during the 3-hour lab session, students are expected to summon the TA for a brief chat about their work. TAs are provided with a list of possible questions to ask at each checkpoint, and a checklist to ensure every pair of students completes all the checkpoints.

The adoption of Real Time Physics as a lab curriculum required the TAs to not only help students develop skills to trouble-shoot apparatus and computer problems [289], but also support inquiry-style learning. However, since our professional development program aims to prepare lab TAs to support principles of guided inquiry and inclusion in student learning, this new curriculum was even more attuned to the TA training provided.

12.3.3 Collecting Qualitative Evidence for TA Buy-In and Growth

As a precondition to being able to evaluate the change in TA behaviors that come as a result of our implementation of lab TA professional development, we sought to determine whether the TAs bought in to the professional development we provided. We relied on two approaches to do this: informal observations and semi-structured interviews. An extensive series of informal observations were conducted by the author and his supervisor, who acted as non-participant observers, during the semesters for which data was collected [72, 228]. These observations occurred during the weekly lab TA meetings, during lab sessions, and occasionally even at times when TAs were preparing for the lab individually. Field notes were recorded, and used as a basis for generative discussion by the author and his supervisor to identify key themes (which then provided focus for future observations) and to evaluate the extent and nature of TA buy-in to the professional development program. These informal observations complemented the structured observations we conducted (see below) and, along with the interviews, were used in part to examine TA buy-in to our instantiation of the lab TA professional development model. In total, more than 200 hours of informal observations were conducted.

We invited 8 TAs to participate in semi-structured interviews about their experiences as a lab TA. The 8 TAs were selected to provide diversity in standing (graduate students in their first and second years of study, as well as those close to graduation), nationality, gender, teaching experience, and our perception of their engagement with the lab TA meetings. The interviews occurred at the end of the semester in which they served as a lab TA, were audio-recorded and transcribed, and lasted 30-45 minutes. All 8 invited TAs participated. The interviews followed a semi-structured format [228]. A list of 'starter questions,' generated by the author and his supervisor, focused on the TAs' perceptions of the structure and goals of the lab TA professional development program; their experiences working with students; how, and to what extent, they used techniques from the lab TA meetings in their labs (such as supporting inquiry learning, maintaining equitable learning, and explicitly talking about the nature of science); and the grading for the lab course. Most participants shared openly and at length about most topics, with little need for prompting. All names are pseudonyms chosen by the interview participants.

12.3.4 Writing Excerpts

In order to understand how TAs' views about learning changed over the semester that they taught, we asked the lab TAs to respond to writing prompts at their first and last lab meetings. The prompts, reproduced in Table 20, were designed to stimulate TAs to write about how they imagined/recalled interacting with students. The written TA responses were read by the author, and all statements describing interactions with students were extracted. The extracts were typically one sentence in length, but in a small number of cases where a single sentence contained multiple potentially-conflicting viewpoints, that sentence was split into two or three phrases.

Table 20: Writing prompts used to generate writing excerpts.

Pre	What can we do as TAs to make sure our students have positive, meaningful, effective learning experiences?
Pre	How will you help students understand how physics concepts connect to the lab, and understand why they're doing what they're doing?
Post	Write about a time when you helped a student learn something in the lab.
Post	Write about an occasion in the lab when you weren't sure what to do at the time. How would you react now?

Next, we sorted the extracts according to whether they indicated that the TA held a transmissionist or constructivist understanding of the nature of learning, using the following definitions, selected from a journal article written for an audience of college educators. In the transmissionist view, the instructor "has the knowledge and transmits that knowledge to the students, who simply memorize the information ... often without even thinking about it. This model of the teaching learning process ... assumes that the student's brain is like an empty container into which the [instructor] pours knowledge" [164]. Meanwhile, for the constructivist view, "...knowledge is a state of understanding and can only exist in the mind of the individual knower; as such, knowledge must be constructed, or reconstructed, by each individual knower through the process of trying to make sense of new information in terms of what that individual already knows... Students use their own existing knowledge and prior experience to help them understand the new material... The [instructor's] role is to facilitate students' interaction with the material and with each other in their knowledge-producing endeavor" [164].

Over three semesters, we collected 61 written excerpts from 24 different TAs, all but two of whom were TAing the lab for the first time. Two physics education researchers first sorted 10 of the excerpts, then compared their results and discussed their sorting process in order to reach a consensus before sorting the remainder of the excerpts. Overall, the sorters achieved an agreement given by Cohen's $\kappa = 0.929$ [54], which is considered "excellent" [97] or "almost-perfect" [175] agreement. Finally, a consensus was reached for the remaining excerpts for which there was disagreement. All 61 excerpts were sorted into one of the two categories.

Trusting this sorting to answer the research question requires the validity of several assertions. First, we claim that the excerpts honestly and accurately depict the TAs' thoughts about student-TA interactions. The TAs who submitted these responses were given time and space to write, assured that their work would be held anonymous, and understood that their supervisor (the lab coordinator) would not evaluate the responses. These conditions, plus the existence of several responses from the end of the semester of professional development meetings that directly oppose explicit instructions from the TA training, suggest that the TAs' responses were honest and true.

Second, we claim that the sorters were correctly reading and interpreting the excerpts. Although there were a few grammatical and spelling errors in the (hand-written) excerpts, they were all clear and coherent. Examples are provided in Table 21. Further, given the high level of agreement between the two sorters, it is highly likely that they were reading and interpreting the excerpts accurately. We note that since the first constructivist statement in Table 21 was from a TA in the traditional lab which focused on "proving formulas", the researchers agreed that it was constructivist for that context.

Table 21: Example excerpts from the TAs classified as Transmissionist and Constructivist.

Category	Example excerpts
Transmissionist	"I will explain the physics model behind their experiment equipment." "I need to map the theoretical background deeply and communicate it in such a way that the students are able to grab the concepts."
Constructivist	"...encourage students to design an experiment by themselves to prove the formula." "I was able to refrain from giving students an answer until they had figured it out themselves... Through questions, I was able to get them to the answer."

The final claim that must be supported in order for this sorting to have validity is that the categories into which the excerpts are sorted must be distinct, and that a TA moving from one category to another is making a meaningful change in their view of student learning. In this case, the two categories of transmissionist and constructivist views are relatively distinct, with a wealth of educational theory supporting this claim [56, 233].

12.3.5 RIOT Observations

The Real-time Instructor Observation Tool (RIOT) is a tool for continuously monitoring and categorizing types of instructor-student interactions [232, 304]. The categories are briefly described in Table 22 and more-thoroughly explained in Refs. [231, 304]. While a wide variety of observation protocols have been developed and used to study instructor-student interactions in labs [235, 299], the continuous recording, ease of use, and applicability of categories made RIOT the best option for this study.

Our RIOT data span three semesters, totaling nearly 200 hours of observations of 24 dif-

Table 22: Descriptions of categories in RIOT relevant to this study.

Category	Description
Open Dialogue	Student is contributing half of the words, actively developing an understanding of physics/lab ideas
Closed Dialogue	Instructor is controlling the conversation, but student is providing some input (asking follow-up questions, answering closed questions, etc.)
Explaining Content (Discussing Concepts)	Instructor is explaining physics concepts while student is a passive recipient
Clarifying Instructions	Instructor is explaining lab procedure while student is a passive recipient
(Actively) Listening	Instructor is actively listening to a student, shown by eye contact, body position, gestures, etc
Active Observing / Watching (one group)	Instructor is paying attention to only one group of students but is not engaging with them in any way
Passive Observing / Scanning (whole class)	Instructor is walking around or on side, not able to pick out individual conversations

ferent TAs. The data were collected in weeks 3, 6, and 11 of the semester. These weeks were chosen to provide initial, midpoint, and end-of-semester data because student enrollment is still in flux during weeks 1-2 and students often skip the experiment during week 12 since they are permitted to drop their lowest lab report grade.

During data collection, the observer sat on a bench at the side of the classroom, using the RIOT app [231] to collect data on his mobile phone. The observer was careful to avoid impacting the regular classroom dynamics. In follow-up interviews with TAs, the TAs confirmed that they did not act differently when the observer was in the room because he was there so frequently.

The validity of RIOT data for answering our research question requires several assertions to be true. First, we claim that the observer could see, hear, interpret, and correctly understand conversations that were happening around the classroom. For example, when the TA is talking with a pair of students on the far side of the room, was the observer able to accurately categorize their interaction based on what he could hear and see? In order to assess this, we video-recorded sample interactions from two different distances and with different audio levels, representing the experience of observing a pair of students working together from across the bench and from across the lab. As part of a regular lunch meeting of the Discipline-Based Science Education Research Center (dB-SERC) at our university, we asked a panel of 27 experienced science educators, to categorize the interactions. At both distances, all 27 educators were able to accurately and correctly distinguish between the Open Dialogue and Closed Dialogue categories of student interaction. These two categories were used because we felt that they were likely the most difficult to distinguish at a distance. The videos we used are available online at [71].

Second, we claim that the observer was accurately and precisely categorizing the TA-student interactions that he observed. To evaluate this, a second observer was given descriptions of the categories, and then simultaneously categorized interactions during four separate 1-hour intervals. Data from one of these intervals is shown in Fig. 13. Agreement between the two observers for the four intervals ranged from Cohen's $\kappa = 0.50$ to 0.73, which corresponds to a "fair to good" [97] or "moderate" to "substantial" [175] level of agreement. This should be viewed as a lower-bound estimate on the agreement between the observers.

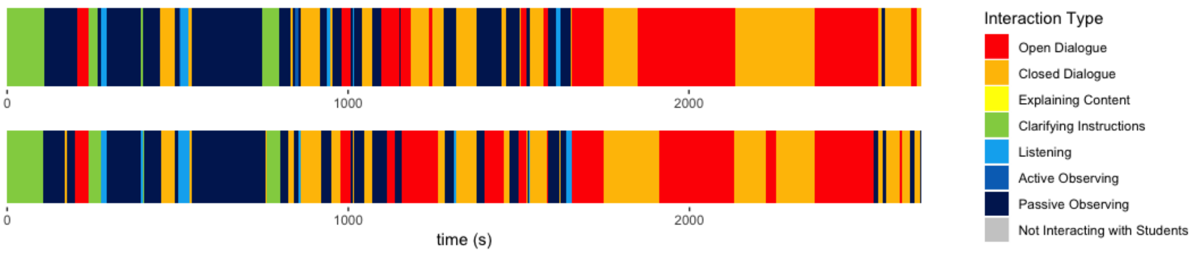

Figure 13: Comparison of two observers' categorizations during a representative 45-minute interval.

In particular, if one observer were pressing the buttons slightly earlier than the other, this systematic temporal shift between two otherwise-in-agreement categorizations would bias the agreement downward [131]. A difference in prior experience in using the RIOT tool and diligence in focusing on interactions may explain some of the other discrepancies between the two observers. We acknowledge that this range of values for Cohen's κ is lower than is reported in other physics education literature, but note that making real-time categorizations as we do here is more error-prone than sorting-type categorization work (as we did with the writing excerpts), for which the researchers have time to carefully reflect before making a categorization decision. Thus, we claim that our categorizations are consistent, and likely accurate, but with relatively sizable random errors. We exercise due caution and restraint in analyzing the data and drawing conclusions that result from this approach.

Finally, in order to draw conclusions from RIOT data, it is necessary that the categories are real and meaningful in the context in which they were observed. For example, were TAs acting in a substantially different way when they are engaging in open-ended dialogue as opposed to closed dialogue? In order to evaluate this claim, we asked our panel of science educators to categorize a variety of video-recorded sample interactions. If the experienced instructors can identify and explain the difference between different interaction types, then those interaction types are likely to be pedagogically meaningful. A panel of 24 educators at dB-SERC were asked to categorize video reenactments of five different TA-student interactions. In all five cases, there was a strong level of agreement among the educators. Fleiss's kappa [97] for these 24 raters categorizing 5 videos using the RIOT categories was calculated to be $\kappa = 0.82$, which is an "almost perfect" [175] level of agreement.

12.4 Results

In order to understand TAs' views about the lab TA professional development program and the extent to which we had established buy-in from them, we conducted interviews with 8 TAs about their experiences as TAs and as participants in the professional development program. Below are summaries from 3 of those TAs. The first two, from interviews with Alan and Emily, are representative and typical of the TAs we interviewed and also of the broader pool of 24 TAs who completed the TA professional development program. Alan and Emily crystalized some factors that were common across many of the other interviews. The third summary is from Ted, a TA with a uniquely low level of buy-in, whose responses help to identify a shortcoming in our professional development program. Ted's responses indicate that for a TA who started with very low interest and self-efficacy, the professional development program was not able to help him develop a higher level of interest and self-efficacy as a facilitator of the inquiry-based labs. Our observations suggest that Ted's experiences were unique among the 24 TAs included in this investigation. All names are pseudonyms.

12.4.1 Alan

A first-year graduate student, Alan reported that his semester as a TA in the introductory lab was his first teaching experience. In the interview, he described how his confidence as a physicist and an educator grew substantially because of his work as a TA: "I was not confident in so many of the materials I had learned before, and I realized that I have to take another look at some of them... How am I going to teach these things some day?"

Alan enjoyed teaching, and his description of a typical interaction with students suggests a well-developed individual interest, on the four-phase framework: "Some people think that the voltage is flowing through the circuit. I saw their answers in their homework, I didn't know they think like this in the lab... You could talk about some things and their confusion. It was very rewarding for me, that I could help them in a small way."

Alan was a typical participant in the lab TA meetings. When asked what he found effective in the professional development, he answered: "There was nothing not effective, all of it was really helpful.. [We] did those sabotage things, [and] introduced some of the common mistakes. I don't know what would happen if we didn't have those sessions." Yet, while he was actively engaged in the meetings, he was sometimes unsuccessful at implementing the teaching strategies in his lab sections. In a discussion about standing on the side of the lab bench and asking questions, rather than touching the apparatus, when students have questions, he confessed, "I have to admit that I wasn't able to do it correctly because sometimes it's really hard to stay away from the thing that they're asking questions about, I have to go there and see what's happening." Here, it appears that Alan's self-efficacy as an educator is developing, but has not fully matured, reminding us that supporting the growth of educators can be a slow process.

For other strategies, such as switching partners to help students avoid developing 'bad habits' like gendered task division [74, 126], Alan implemented the strategy introduced in the TA professional development, but clearly kept thinking about it, as he explains: "When I ask a question in the check-in, most of the time one of them answers... If I were to do everything again, I would change their partners every session, not once a semester. I really notice that some of them keep staying away from the experiment, and their partner is doing all of the things. It's one of the most important things that we have to work on." Whether this technique would be effective is perhaps less important than what it says about Alan's well-developed individual interest and engagement as a TA.

12.4.2 Emily

Emily was in her fourth year as a graduate student, doing research in experimental condensed matter physics. This was her first semester as a TA. Like Alan, she rapidly developed an individual interest as a lab TA. Although initially planning to pursue a career in industry after graduating, by the end of the semester she was considering a teaching role instead: “I went from, a year ago, absolutely, post-doc or private researcher. And now, I’m thinking of ... teaching full time. This [lab] has done a lot to show me that I can interact with students and have a lot of fun.”

Emily was enthusiastic about the lab TA meetings. Asked whether they helped her prepare for her work as a lab TA, she replied, “Definitely yes, the meetings helped, [especially] the parts of the meetings where we go see what the set-up is like, and go practice it.” Emily found the sabotage activity to be especially useful. She also appreciated having the opportunity to reflect and connect with fellow lab TAs, reflecting growth in her self-efficacy while noting that “it’s nice knowing that we can all check in with each other, just knowing that the time is there.” Emily reported that the nature of science activity was less effective, and suggested that discussions about the nature of science might be better positioned as questions for the lab report rather than part of the discussions between TAs and students.

Unlike Alan, Emily told us that the strategies to support inquiry learning “came really easily.” She relished the guide-on-the-side role, and preferred to help students find their own answers by asking them questions. “I loved the check-in after Investigation 2, capacitors in series and in parallel, trying to get them to understand why they add up the way they do. Some students had the answer, and some didn’t. It was really fun helping them arrive at that, because it was so intuitive and physical. I enjoyed that part the most, when they let me ask them a bunch of questions.” Like Alan, Emily’s responses suggest a lab TA with well-developed individual interest and positive self-efficacy growth as an educator in her work by the end of the lab professional development program.

12.4.3 Ted

Ted was the lab TA who was least willing to engage with the activities during our lab TA meetings. A fifth year graduate student, Ted studies theoretical condensed matter physics. Unlike Alan and Emily, who demonstrated buy-in to the principles of supporting inquiry and inclusive learning, among others, Ted remained skeptical about the value of learning pedagogical techniques in the lab TA meetings. Asked if he found the meetings valuable, he replied, “A bit. I think the best way to hold the lab meeting is to give us a complete instruction about the lab. For example, we could do the experiment together." Ted saw the lab meetings as a place to learn about the lab procedure and apparatus, rather than an opportunity to learn new teaching strategies and get practice interacting with students.

While Ted saw the pedagogy, he did not adopt it. Asked if he sought to help students to frame their own questions, he answered, "There were too many students. I can't just instruct them step by step. There's not enough time." Similarly, he decided not to adopt the strategy of talking to quieter students in order to counteract conversational dominance, saying, "Trying to ask the student who is not good at physics, I think they cannot answer your question."

However, while Ted did not buy in to the teaching approaches we shared during the professional development program, he did recognize the problems those approaches were designed to address. He described a case of gendered task division, saying, "One student is very good at physics, they will try to do most of the work." However, while he was concerned about this happening, he felt unprepared to address it. When asked how he responded when such situations arose in his lab, Ted said, "I have no very good ideas. I told them to mix their groups several times, but it didn't help."

This situation of a TA who sees the need for certain teaching strategies, but does not use them when they are presented to him, suggests that Ted did not consider that the teaching strategies he was being presented had merit. In other words, he did not buy in to these strategies. Moveover, and exceptionally, during his 30-minute interview Ted told no stories about interactions with particular students, expressed no curiosity or interest in a student's struggle or success, and indicated no pride or interest in his teaching. Ted's lack of buy-in could be understood by hypothesizing that Ted never felt a triggered situational interest in helping students learn during his work as a lab TA.

Aside from Ted, however, we may broadly claim that our instantiation of the model of lab TA professional development had a positive impact on TAs. Not only did most of the TAs (except for TA) participate actively in the lab TA meetings, but they continued to reflect on their role as a TA outside of the meetings and worked to improve the way that they supported student learning in their lab sections. Thus, since most of the TAs have demonstrated that they bought in to our instantiation of lab TA professional development,

Table 23: Categorization of TA writing excerpts into two views of teaching.

	Pre	Post
Transmissionist	16	9
Constructivist	6	11

we may consider the evaluation of our implementation of that instantiation.

12.4.4 Did TA Views About Teaching and Learning Change to be More Supportive of Inquiry Learning?

Over three semesters, a total of 61 writing excerpts were collected: 34 from 24 TAs at the start of the semester, and 27 from 21 TAs at the end of the semester. Representative example excerpts are presented in Table 21. However, it is the TAs – and not the excerpts – that we wish to compare. If we counted only the excerpts, one TA whose writing was the basis of three excerpts would be over-represented, and perhaps introduce bias in the results. Thus, the 61 excerpts were associated with the TAs who wrote them, and these data are indicated in columns 2 and 3 of Table 23. Since the writing was submitted anonymously, it is not possible to look at the change in views for individual TAs. Three TAs are excluded from this table (2 pre and 1 post) because their writing resulted in two or more excerpts that contained both transmissionist and constructivist views of teaching. The Fisher-Irwin test [97] allows us to reject the hypothesis of independence between TA view of teaching and the pre/post condition ($p < 0.05$), indicating a positive and statistically significant effect when comparing the 22 (pre) and 20 (post) TAs whose writing provided unambiguous transmissionist or constructivist views of teaching.

12.4.5 Did TA Behaviors Change to the More Supportive of Inquiry Learning?

RIOT observations for three semesters are shown in Fig. 14. The fraction of lab time is shown, with colors indicating the different TA-student interaction types. Each week (i.e., each column) is the mean of these fractions over 7 or 8 TAs, depending on how many were observed that week. Each semester, a new batch of TAs was observed, so these changes indicate how TAs' interactions with students change during their first semester working as a lab TA. For clarity, no error bars are displayed. Typical values of the standard error in the mean are 2-3%.

During the first semester of observations, the lab was run using a highly-structured lab curriculum [50] and no specialized lab TA professional development was provided. The shift from watching and scanning toward closed dialogue (i.e., from observing to talking) aligns with our observations that the TAs became more confident in talking with students over the course of the semester. However, the TAs spent very little time engaging in (open-ended) dialogue that allowed students to actively develop understanding or expertise.

In the second semester, the highly-structured curriculum was retained while specialized lab TA professional development was introduced via weekly, one-hour lab TA meetings. While the distribution of TA-student interactions is comparable to those in the first semester during Week 3, by Week 6 the difference between the two semesters has become clear and distinct. Unlike in the first and third semesters, in the second semester the behaviors of this new batch of TAs continued to evolve during the semester, perhaps reflecting slow progress as the TAs negotiated tension between the traditional, highly-structured lab learning activities and the encouragement from the lab TA professional development program that TAs support student learning using techniques designed for inquiry learning.

In this second semesters, the TAs spent substantially more time actively listening to their students and engaging in open-ended discussion (i.e.: as in the first semester, TAs went from observing to talking, but now with more open dialogue). In our observations, we noted that the TAs seemed more comfortable in their roles as educators than they had in past semesters. In general, the TAs were more willing to simply listen to their students, rather than always feeling the need to provide answers. This semester in particular showcased the

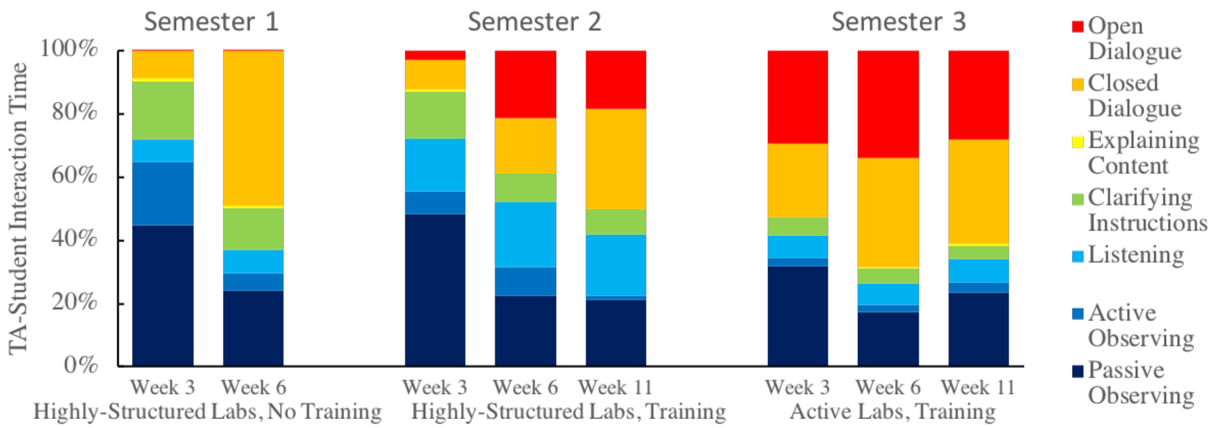

Figure 14: Comparison of TA-student interactions across three semesters, averaged over all observed TAs.

utility of the lab TA professional development program, as the training helped TAs engage in epistemologically beneficial instructional practices even though the traditional, highly-structured lab curriculum used during this semester was built around learning activities that were not designed to develop students' epistemology.

The third semester saw a switch to a Real Time Physics-based active learning curriculum [278]. This switch included adopting checkpoints, as described previously, and continuing to provide professional development activities in the lab TA meetings. The difference between the second semester and the third semester can be attributed to the change in pedagogy. In this semester, nearly a third of TA-student interaction time was devoted to open-ended dialogue, including as early as week 3, and a similar fraction to closed dialogue. We observed that the TAs this semester generally demonstrated both confidence and purposefulness when interacting with students, and were especially likely to employ the strategies for supporting inquiry learning that were introduced as part of the lab TA professional development. In other words, the benefits of the professional development program on TA practices become evident earlier in the semester in the third semester in which inquiry learning learning was used compared to second semester when the lab was traditional.

12.5 Discussion and Summary

The findings described previously suggest that, on average, our professional development program helped improve lab TAs' views about teaching and learning and their concomitant teaching behaviors in the lab. Our instantiation of the lab TA professional development program, designed using the framework of the cognitive apprenticeship and expectancy-value theories, did an adequate job of moving many graduate student TAs toward a higher level of teaching effectiveness. However, as illustrated by the interviews with two representative TAs and one TA with a uniquely low level of interest and self-efficacy, the average results presented here mask inadequacy of certain aspects of the implementation of effective pedagogical practices. For example, Alan struggled to avoid touching the students' apparatus and Ted largely did not buy in.

As viewed alongside the measuring-stick of the four-phase framework of interest development, many TAs (including Alan and Emily) experienced substantial growth in their interest. One TA (Ted), on the other hand, did not engage deeply during the activities which were part of the lab TA professional development program, nor did he employ the teaching strategies in his lab. It appears from the interviews that this TA, about a year away from defending his Ph.D. thesis and the only member of his cohort in the pool of lab TAs in that semester, may have viewed the lab teaching role as an unwanted burden. In order to account for Ted's lack of involvement, we refer back to expectancy-value theory. Unlike most of his fellow TAs, Ted came with a very low level of interest and did not find enjoyment in his work as a lab TA throughout the semester. Thus, our focus during the professional development program on developing interest with inspiration from the four-phase interest development theory was not appropriate for a TA with a very low level of initial interest as an educator. Instead, in order to adopt the pedagogical techniques and approaches advocated in our professional development program, Ted may have required that the program be built around a utility value model [145] focusing, e.g., on more explicit discussions of how developing teaching skills could help his career goals, or on how improved teaching practices might improve his course evaluations.

Both the excerpt categorization and the RIOT observations indicated that, while improved in their capacity to support inquiry learning, the instructional views and practices of the TAs continued to be an admixture of transmissionist and constructivist views. For example, in the 'post' condition, 9 TAs continued to describe student learning using a transmissionist view, and around 40% of TA-student interactions in Week 11 with the active labs and TA training continued to consist of closed dialogue. We interpret these results as indicative of two things. First, the process of developing expertise as an educator should be viewed as a long and sometimes difficult process: those 9 TAs who continue to ascribe to a transmissionist view of teaching at least in some cases may have shifted in their thinking but still have some distance to go before they fully adopt a constructivist view. These TAs may benefit from more practice teaching, and perhaps another semester as a lab TA with our professional development program. Second, we note that masterful educators typically use a mix of methods. In our labs, TAs are given enough freedom and responsibility that it would be remiss to only engage in open dialogue with students, or to never clarify the instructions for them. Thus, the argument we seek to make is that the overall balance of TA-student interactions was better-attuned to supportive inquiry learning with the professional development than without it.

Equity was an important consideration in the design and implementation of our lab TA professional development program. Several of the lab TA meetings included activities that sought to help TAs learn to recognize and combat inequitable student work, such as gendered task division and conversational dominance. As was the case for Alan in the interview described previously, it was common for TAs to engage with these activities, and to think about how they could respond to inequities. In practice, our informal observations suggested that most TAs adopted techniques for responding to conversational dominance (such as posing questions directly to the quieter partner) and were diligent about rotating group composition at least once during the semester in order to reduce the impact of gendered task division. The higher level of engagement with these activities suggests that most TAs bought in to the topic of equity although more work needs to be done.

In summary, our model of lab TA professional development based on weekly lab TA meetings took into account both cognitive and motivational aspects to engage TAs effectively. This type of lab TA professional development can be successful at preparing graduate student

TAs to effectively support student learning in the labs. Moreover, our findings suggest that over the span of one semester, TAs who engage in lab TA professional development can develop effective attitudes and approaches to supporting student learning. These results hold regardless of the nature of the curriculum used in the lab course. Our lab TAs demonstrated the use of effective forms of student-instructor interactions regardless of whether the students were in a lab course that used a traditional, highly-structured curriculum or a transformed, inquiry-based approach.

Equally importantly, the success of lab TA professional development at achieving these goals is contingent at least partly on the buy-in and engagement from the graduate students who undertake it. Including expectancy-value theory and the 4-phase interest development theory as the motivational framework inspiring our work was important for success of the program and ensuring buy-in and engagement from TAs. We also note that the close relationship between the lab TA meetings and TAs' instructional work in the labs may have allowed a synergy that is not available in all cases of TA professional development.

Several avenues for future work are suggested by the interviews. For example, as Emily pointed out, discussions about the nature of science are difficult for TAs to initiate. Based on these types of factors, our lab curriculum has been revised to incorporate explicit discussion of this important topic in the questions that students answer during the lab, rather than asking TAs to initiate conversations about the nature of science. Another place for improvement was highlighted in the interviews with TAs such as Alan and Ted, who described a need to continue developing and refining strategies to ensure equity and inclusion of all students in these types of physics labs. The task division and dominance activities have been useful first steps, but we are also committed to refining the lab curriculum and environment to eliminate unequal opportunities for student learning in these types of labs.

13.0 Student Reflections on Online Learning

Over the past year, colleges and universities around the world have pivoted to remote instruction as a way to provide educational continuity for learners who are able to attend remote classes during the COVID-19 pandemic disruption. Early reports focused on the impact on teaching and learning that arose from the transition to emergency remote instruction in the northern hemisphere spring of 2020 [43, 68, 105, 168, 310]. The switch to remote instruction in the spring of 2020 was hasty, giving instructors no time to develop or cultivate online resources for students, design effective learning experiences, or address the specific affordances and constraints associated with online learning.

However, by the start of the fall semester, many instructors made heroic efforts to make themselves acquainted with the challenges of supporting student learning online, the capabilities of modern technology in delivering remote instruction, and resources and curricular materials available to support instruction [32, 42, 67, 227, 247].

Many instructors worked extremely hard and sought to deliver thoughtfully-designed online courses for the first time in the fall of 2020. Given the current, ongoing, need for remote instruction because of the pandemic, and anticipating that some colleges and universities may increase the fraction of courses they offer online in coming years, it may be productive to reflect on what worked and what didn't from students' perspective for online classes in the fall of 2020.

At our institution, a large research university in the USA, fall 2020 instruction was fully remote at the start and end of the semester, with a period during the middle when instructors were able to conduct their classes in a hybrid format if they wished. Few instructors employed hybrid instruction during the weeks when it was possible and, for those who did, only a very small number of students attended in-person. In this chapter, we analyze survey responses from 1400 students in physics courses at all levels (which is part of a pre/post survey our department administers for departmental assessment) and interviews with 37 physics students in fall 2020 in order to understand students' perceptions of the affordances and limitations of online college physics instruction at our institution. The surveys were given to students in 12 physics courses with 10 different instructors, with an overall response rate of 80% and response rates from individual classes ranging from 63% to 100% during the Fall 2020 semester. Student responses shed light on five areas of instruction: lecture-based classes, active learning classes, lab classes, community and collaboration, and assessment.

As we collected and analyzed these data, we recognize the ethics of teaching and reflecting on student learning during the pandemic. Many of our students have been severely adversely affected by the pandemic in many ways. In fact, the pandemic has brought out the inequities in higher education clearly. Many students have had health issues or have lost loved ones, and others have faced challenging circumstances or paused their studies. We seek to be conscientious about the needs of students who are not represented in our data due to hardships caused by the pandemic, about the additional serious challenges facing students who continued their studies, and about inequities related to who was able to continue their studies. We focus on students' perceptions of online learning rather than potentially biased measures of instructional effectiveness. Our survey data were collected as part of a routine departmental assessment that we conduct every semester. The interviews were conducted with volunteers who were paid for their time in an IRB-approved study. The data collection and analysis were originally conducted to provide guidance to instructors in our department on how to tune their online physics courses for the Spring 2021 semester, and this analysis was done only after that task was complete.

13.1 Lecture-Based Classes

In interviews and survey responses, students expressed both satisfaction and frustration with how the lecture portion of their science courses had been converted into an online format. Among the positives, students described practical advantages to attending lectures online. They could see the slides or whiteboard better and hear the instructor more clearly. Studying from the comfort of home was a benefit, as well. Students described being better able to maintain a restful sleep schedule because they didn't need to wake up early to commute to class, feeling more physically comfortable studying from the furniture in their homes, and feeling less anxious because they weren't (hyper)visible [239] to their classmates. They reported being more likely to attend class, and more likely to be on time when they did.

The biggest advantage to live online lectures, according to students, was that they could re-watch the lectures (and rewind some parts several times to understand concepts) in order to catch up on the concepts they missed, refresh on ideas later in the semester, or fill in gaps in their notes. In response to a survey question that asked students to identify the most positive outcome of remote instruction, 21% noted the ability to re-watch live lectures. For example, one student appreciated "the ability to [go] back and look through the recorded lectures when I'm having difficulty understanding the concepts or need future clarification."

However, the practical benefits of online lectures were balanced with substantial drawbacks. The most significant effect described by students in our survey was a decrease in their focus and motivation because their class was online. While home is comfortable, it is also full of distractions. One student noted, "I definitely pay attention less than when I was in a normal in-person class, so I get less out of each lesson." Another student pointed to "the issue of staying on task. It's difficult being in your room and staying focused on lectures or homework. It's very easy to get sidetracked on your cell phone or by cleaning your room." Without social cues to pay attention in class, it was easy for students to lose focus. A third student noted, "I could never focus on anything I was doing. Being stuck inside and unable to see how my peers worked made me less motivated." The feelings of loss of focus and demotivation went hand-in-hand with poor mental health, with students specifically describing feelings of isolation, loneliness, burn-out, and 'Zoom fatigue' [15]. In response to a question that asked students to identify the most negative outcome of remote instruction, three of the seven most common themes were student concerns about motivation, focus, and their mental health.

While many students articulated both positive and negative aspects of online courses, first-year students in particular seemed negatively impacted by online learning. Since they were new to the university, they found the online environment made it much harder to adapt to college science learning. Reflecting on a first semester of college spent taking online courses, one student described the experience as feeling "fake." The transition from high school to college is not always easy. It is possible that the affordances of in-person instruction provided a bridge to help some students through the transition, such as by providing opportunities for students to collaborate with one another. With classes moved online, students no longer benefited from some of the informal norm-setting and expectation balancing that may happen when students talk with each other before, during, and after class. It might be productive for colleges and universities to think about ways they could provide a more structured transition from high school to college style learning.

First-year students found online lecture courses to be “not real”, ineffective and disappointing, while students in their second year and beyond described a balance of advantages and disadvantages to studying their lecture-based science courses online. The technical benefits of online lectures such as being able to hear and see more clearly, not needing to commute, and being able to re-watch lectures were balanced by concerns about motivation, focus, and mental health.

13.2 Active Learning

The most common alternatives to traditional (but online) lecture-based courses at our institution were flipped classes [163]. In flipped classes, students were expected to watch lecture videos before attending class, where they would mostly engage in collaborative skills practice activities with coaching from the instructor. In some classes, the instructor used the Zoom polling feature or Top Hat to replace clicker questions, with grade incentive provided for answering questions. In some cases, students were also graded for the correctness of their polling answers. Breakout rooms were used for group discussions.

However, in many classes, students were given no grade incentive for completing the in-class collaborative activities or out-of-class homework, and solutions to these activities were posted online. In such classes, students reported that the lack of grade incentive meant that they simply stopped doing these activities on time and prioritized other classes. Moreover, a negative feedback loop occurred such that some students who attended class were unprepared for collaborative work, deflating the quality of group skills practice activities, which students quickly began to consider “useless.” One student described, “working on a worksheet then brings back the problem ... where people don’t want to participate as much. So then, at the end of the day, you’re not really getting group activity help.” Students in classes where weekly assignments did not have an associated grade incentive attempted to ‘cram’ before mid-term and final exams by watching videos and rushing through practice problems and browsing over their posted solutions, but this strategy was largely ineffective. However, in classes where weekly assignments and collaborative work was graded, students continued to participate throughout the semester and felt that they had learned well and were better-prepared for their exams.

In some classes, the instructor sought to provide completely asynchronous instruction for students, re-purposing scheduled classes into drop-in office hour sessions. The students found it very difficult to learn in these classes, as the lack of synchronous interaction decreased their capacity to learn and participate in the class. One student commented, “it was harder to make sure that you were on pace.” Without incentive to keep up with lecture videos and homework, most students skipped office hours and quickly fell behind. A useful comparison can be made to asynchronous Massive Open Online Courses (MOOCs), where typically only a few highly driven students complete the course since there is little motivation from grades or interactions with the instructor or peers to watch the lectures or complete assignments or assessments regularly [271].

Thus, when weekly activities were not graded, students in flipped and asynchronous online classes struggled to keep up. Students struggled to self-regulate [320] when they did not encounter collaborative work or an explicit grade incentive. It may be wise for instructors to ensure that their online classes include a synchronous component, and that students are incentivized to complete homework including watching of videos, collaborative work, and other activities during class by assigning a small grade incentive to this work. Of course, instructors should be flexible with students who are unable to participate in these activities each week due to challenging situations and grade them using other means.

13.3 Lab Courses

We asked the students in our interviews to describe both their physics labs as well as other labs they were taking. One common approach [98] was for instructors to video-record experiments and have students conduct analysis from the data collected in the video-recording. While they appreciated the effort that goes into making such videos and the paucity of alternatives, the students we interviewed were skeptical of this approach, which they felt served to simplify the labs too much and took away the opportunity to participate in hands-on science. One student commented, in reference to this approach to doing labs, that "you don't really do much work and you don't do much thinking."

Another popular approach to doing labs online is to use simulations [98]. Although simulations also simplify lab-work, reducing instruments to cartoons, they still provide students opportunities to make decisions and collect their own data. One student commented, "I really enjoy ... doing all these online simulations. Everything has been so helpful in truly, truly understanding the material to the best of our ability."

A smaller number of labs found ways for students to collect their own data, either by having students pick up apparatus, rotate through labs one at a time when it was possible to do so, or use household goods as apparatus. While students appreciated the opportunity to do their own experimentation, none of these approaches worked for everyone. Some had difficulty tracking down particular supplies. Some students had difficulty setting up and operating apparatus without in-person support from peers or the instructor. Others found that without that support, the amount of time they spent troubleshooting their apparatus ballooned.

Students in our introductory physics labs used a combination of simulations and hands-on experiments at home using the IOLab system. However, regardless of the type of lab-work that students conducted, they felt that collaborative work was an essential part of the lab experience. Students commented that labs that did not include collaborative work felt incomplete. Reflecting on labs that did include collaboration, one student commented, "I think collaborative activities with your lab group is very important. I think it's very important that we still do them." Many students also missed the in-person aspect of the labs in which they shared the experimentation with other students in a shared space.

13.4 Collaboration and Communication

13.4.1 Community and Studying Outside of Class

Most of the students we heard from via surveys and interviews expressed a desire to collaborate with one another, both in the classroom and out of the classroom in study groups, but faced difficulties doing so. When asked about the most negative outcome of online instruction, the most common response described a lack of community engagement and group work. This inhibited students from learning from one another, with one student stating, “I found it difficult to find help because I do not know how to interact with people I haven’t met or know what they look like”. The lack of collaboration impacted their learning of the material as well as their motivation. It was harder for students to form natural connections with their classmates, e.g., with another student, mentioning, “I miss being able to discuss physics problems and exams with classmates while walking back from class.”

Although it was harder for students to work together during the semester, they found benefits from collaborating with their peers when they were able. For example when asked about the most positive part of online classes, one student mentioned that, “working on physics problems in a group made me enjoy physics more and really helped me learn”. Since many students weren’t able to work together in-person, they found new ways to help each other and provide some of the peer support they missed because of online instruction. Another student stated that one benefit of remote instruction was “having more online groups for each of my classes where students support and motivate each other”. It could be valuable in future in-person courses for instructors to provide an online environment in which students are given tools and formats to support one another.

13.4.2 Breakout Rooms

Some professors encouraged collaboration in their classes by having the students work together in video-conference breakout rooms. Students described how the structure and group composition of breakout rooms played an important role in determining its success, with some students describing how other members of their group did not participate in breakout room discussions. One student noted that, "I think over Zoom it did make it a little difficult because nobody really wants to turn their camera on and talk in general." Another student talked about the difficulties of working together over Zoom, stating, "even in recitation, it was difficult to work on assignments together. We all pretty much did our own work and barely talked problems out." Therefore, students ended up working alone on assignments that were supposed to involve collaboration.

Students also talked about situations in breakout rooms where group work was successful. Some students mentioned that a grade incentive for group work was an important factor. For one student in the physics lab, the activities were successful "because the collaborative activities were mandatory to do and they were graded. So people came to them and worked on them." Other students were able to have successful conversations in breakout rooms if they needed to share an answer with the class when they were done. Structure and incentivized preparation for breakout sessions seem to have been essential for science courses, but less important for non-science courses in which students felt they could participate in conversations even if they hadn't completed the preparation. One student described breakout rooms in a world religions class as more talkative and easy to participate in without having done the class preparation for this reason. Then, since the breakout room had productive discussions, the instructor required students to share with the class afterward. The student explained, "we actually have to come up with an answer and then someone has to share it. I feel like the fear of getting called on and not having an answer prepared makes people talk a little bit more."

Another advantage of breakout rooms is that there was no noise from other groups talking at the same time. Another disadvantage of breakout rooms is that it is difficult for the instructor to know which group needs "nudging" or help unlike an in-person class where it is easy for the instructors to notice the groups that are productively engaged in discussion and the ones that are not.

Along with a grade incentive, teaching assistants (TAs), undergraduate teaching assistants (UTAs), or undergraduate Learning Assistants (LAs) could play a valuable role in stimulating small group discussions. One student explained, "I think a big motivator is when the TA does come around to the breakout rooms for people to speak up." TAs might be trained and provided with a 'script' of example questions they could ask to help steer the discussion in a productive way. When they were structured, incentivized, and supported effectively, students typically reported that breakout rooms were a valuable and productive aspect of their online learning.

13.4.3 Back Channel Chat

Some instructors were able to use the chat feature in Zoom to provide students with a back channel to ask questions during class. In some large classes with an assigned TA, the instructor would task the TA with monitoring the chat, responding to questions when possible, and flagging some questions for the instructor to answer. It can be challenging for students to raise their hands and ask questions in class, knowing that by speaking up they are subjecting themselves to the judgment of their peers. The back channel reduces this fear. One student noted, "I don't feel shy to ask questions because I'm not in-person." However, in some classes, the back channel became dominated by a small number of personalities, and some students reported being so engrossed in the discussion on the chat that they were no longer paying attention to the lecture. Productive uses of the chat as a back channel typically required the establishment of norms for the use of the chat. When in-person classes resume, it may be worthwhile exploring how in-person lectures could incorporate recording and back channels to support student learning. It may be especially valuable to have a TA monitor the chat, if possible.

13.4.4 Office Hours and Communication

Most students found that online office hours were easier to attend. In addition to not needing to walk over to the physics department, some students also appreciated that they could log in to office hours with their camera off and listen to other students' questions if they weren't yet ready to ask their own questions. However, other students felt that online office hours were more difficult to attend. For example, one student noted that it was "harder to 'pop by' the office for a quick question." In some cases, instructors asked students to contact them to schedule an office hour visit. This approach seemed to deter students from seeking extra help, as it made the office hour feel more formal and removed the possibility for students to listen in with their cameras off until they felt comfortable participating actively.

Several students described finding it more difficult to ask questions or get feedback from their instructors, and difficulty in communicating with their instructors was a common theme in the survey responses. Some students commented that their instructors didn't always reply to their emails. According to one student, "I completely rely on my professors' validation and comments to know how well I'm doing in this class. I couldn't get any of that because I couldn't interact with my professors." Virtual drop-in tutoring (via Discord) was perceived to be less effective than in-person tutoring had been. Some other students struggled to understand the structure or requirements of their courses, especially when multiple online tools were used. One student disliked "checking multiple platforms for assignments and potentially missing assignments because of it."

The question of whether students should have their cameras on or off came up frequently in our interviews and survey responses. While students typically understood that instructors preferred to lecture to classes that had their cameras on, they also expressed a variety of reasons why they sometimes preferred to keep their cameras off. These included concerns about internet bandwidth, appearing on-camera in their pajamas, and other issues related to attending class from home. However, most students preferred that their peers turn their cameras on for breakout room discussions, and expressed frustration when their classmates did not do this. It might be productive for instructors to set clear expectations for camera usage in their classes, such as strongly recommending but not requiring their use in breakout rooms.

13.5 Assessment

Students expressed a wide variety of opinions about assessment in online courses. Some found online quizzes to be easier, online tests written at home to be less stressful, and open-note exams to provide a better opportunity to demonstrate their understanding of physics concepts. Others disliked short, timed quizzes; struggled with scanning and uploading work; and felt that the assessments were more challenging than they had been in-person. Some students expressed enthusiasm for video-based assessment that either required them to record short videos in which they solved problems or that involved a short, low-stakes oral examination via Zoom.

Overall, the students preferred classes that adopted a strategy of frequent, low-stakes assessments. Frequent assessments provide plenty of feedback, low stakes keep anxiety low, and the flexibility inherent in frequent low-stakes assessments can be an important affordance for students. By decreasing anxiety, frequent low-stakes assessments may make introductory STEM courses more equitable [16, 57].

One physics instructor implemented a group portion to the exams in their class [150, 307]. In this two-stage group exam, students worked in groups of four on typical physics problems. Then, the next class, the students individually wrote responses to questions that asked about the strategies and physics concepts they used to solve the problems. The instructor's goal was to provide students with an opportunity to collaborate with peers during the assessment, while still providing a measure of individual accountability, as a way to decrease the pressure students might feel to engage in academic misconduct during assessments. One student shared their experience with the two-stage group exams, stating, "I felt that working with my group was rather enjoyable and it helped ease some stress I had had about physics in the past." Other instructors divided the total points across the semester so that even though the final week's assessment was cumulative, it was not worth as many points as it typically is. Students generally appreciated this approach, in which one exam did not count for too much of their grade. Instructors should consider some of these approaches in their in-person classes as well.

13.6 Discussion and Conclusion

The fall 2020 semester was the first time that many experienced instructors planned and delivered online instruction. Therefore, it is useful to reflect on the lessons we learned during this semester in order to improve how we design and conduct online classes in the future.

In this chapter, we summarized student perceptions related to online lecture classes, active learning classes, labs, collaboration and communication, and assessment. Students described both positive and negative aspects of lecture-based classes, including the benefits of re-watching lectures as well as decreased focus and motivation. Flipped classes that included a low-stakes grade incentive for weekly activities and synchronous work were effective, while those without grade incentive were not. A variety of approaches to lab-work were described by our students, with the importance of opportunities to collaborate being especially important. Breakout rooms, chat, and office hours were viewed as being potentially valuable, although there was always a need for instructors to be mindful about structure and norms to make them effective. Assessment strategies that included frequent, flexible, low-stakes assessments were preferred by students.

One further issue that came up frequently in survey responses and interviews was the instructors' tech savvy. Many instructors were applauded for introducing new digital tools or, even better, for using standard online tools effectively. In some cases, however, students felt that considerable instructional time was lost by instructors who struggled with their technology. Considering how the instructors learned to use these tools in a short time, some difficulty with technology usage was inevitable.

We encourage instructors to attend to students' perceptions about online learning when designing learning for future online classes.

14.0 Future Directions

I am hopeful that the results and conclusions presented in previous chapters may prove useful for instructors who are interested in improving equity and inclusion in their introductory physics labs. However, while I have addressed mechanisms in the lab that may be responsible for inequities, more work is needed.

There is a need for the development and evaluation of lab curricula that fully incorporate the principles described in previous chapters, as well as those studied and reported elsewhere. While I propose that seeking to reduce the impact of gendered task division, for example, could have positive impacts for all students, it is not yet clear that this works in practice.

Since much of this analysis has focused on the issue of gender inequity, there is the need to consider how inequities in the lab can lead to barriers to inclusion for other aspects of identity. Chapter 6 touched briefly on race/ethnicity, with results suggesting what happens with gender may not happen with race/ethnicity due to, among other things, the masculine nature of physics. Non-visible aspects of identity such as sexuality, ability status, and class may also be important factors, and might be important to consider in future analyses.

The analyses reported here come from one research-focused university, in which students from historically underrepresented minority groups in physics are minorities in the student population. It is not clear which of the dynamics and mechanisms reported in previous chapters would remain relevant at universities in other parts of the U.S.A. or elsewhere in the world, at two-year colleges or liberal arts colleges, at historically Black colleges and universities or at Hispanic-serving institutions or at traditional women's colleges.

While there is work still to do, I am hopeful that research-informed transformations can improve equity in introductory physics labs, removing barriers and improving inclusion in the discipline of physics for coming generations.

Bibliography

[1] AAPT Committee on Laboratories. *AAPT Recommendations for the Undergraduate Physics Laboratory Curriculum.* AAPT, 2015.

[2] Nina Abramzon, Patrice Benson, Edmund Bertschinger, Susan Blessing, Geraldine L. Cochran, Anne Cox, Beth Cunningham, Jessica Galbraith-Frew, Jolene Johnson, Leslie Kerby, Elaine Lalanne, Christine O'Donnell, Sara Petty, Sujatha Sampath, Susan Seestrom, Chandralekha Singh, Cherrill Spencer, Kathryne Sparks Woodle, and Sherry Yennello. Women in physics in the United States: recruitment and retention. *AIP Conference Proceedings*, 1697(1):060045, 2015.

[3] D Ahrensmeir, JMKC Donev, RB Hicks, AA Louro, L Sangalli, RB Stafford, and RI Thompson. Labatorials at the university of calgary: In pursuit of effective small group instruction within large registration physics service courses. *Phys. in Canada*, 65(4):214–216, 2009.

[4] Saalih Allie, Mogamat Noor Armien, Nicolette Burgoyne, Jennifer M. Case, Brandon I. Collier-Reed, Tracy S. Craig, Andrew Deacon, Duncan M. Fraser, Zulpha Geyer, Cecilia Jacobs, Jeff Jawitz, Bruce Kloot, Linda Kotta, Genevieve Langdon, Kate le Roux, Delia Marshall, Disaapele Mogashana, Corrinne Shaw, Gillian Sheridan, and Nicolette Wolmarans. Learning as acquiring a discursive identity through participation in a community: improving student learning in engineering education. *Eur. J. Eng. Educ.*, 34(4):359–367, 2009.

[5] Brian J Alters. Counseling physics students: A research basis. *The Physics Teacher*, 33(6):413–415, 1995.

[6] Susan A Ambrose, Michael W Bridges, Michele DiPietro, Marsha C Lovett, and Marie K Norman. *How Learning Works: Seven Research-Based Principles for Smart Teaching.* John Wiley & Sons, San Francisco, CA, 2010.

[7] American Association for the Advancement of Science et al. *Science for all Americans.* 1990.

[8] Katherine Ansell and Mats Selen. Student attitudes in a new hybrid design-based introductory physics laboratory. In *Physics Education Research Conference 2016*, PER Conference, pages 36–39, Sacramento, CA, July 20-21 2016.

[9] Angelo Armenti and Gerald F. Wheeler. Hawthorne effect and quality teaching: Training graduate teaching assistants to teach. *American Journal of Physics*, 46(2):121–124, 1978.

[10] Arnold B. Arons. Guiding insight and inquiry in the introductory physics laboratory. *The Physics Teacher*, 31(5):278–282, 1993.

[11] Reuben S. Aspden, Miles J. Padgett, and Gabriel C. Spalding. Video recording true single-photon double-slit interference. *American Journal of Physics*, 84(9):671–677, 2016.

[12] National Science Teachers Association et al. Nsta position statement: The nature of science. *Document retrieved*, 3(18):03, 2000.

[13] Lauren M. Aycock, Zahra Hazari, Eric Brewe, Kathryn B. H. Clancy, Theodore Hodapp, and Renee Michelle Goertzen. Sexual harassment reported by undergraduate female physicists. *Phys. Rev. Phys. Educ. Res.*, 15:010121, Apr 2019.

[14] Margarita Azmitia and Ryan Montgomery. Friendship, transactive dialogues, and the development of scientific reasoning. *Social Development*, 2(3):202–221, 1993.

[15] Jeremy N Bailenson. Nonverbal overload: A theoretical argument for the causes of zoom fatigue. *Technology, Mind, and Behavior*, 2(1), 2021.

[16] Cissy J Ballen, Shima Salehi, and Sehoya Cotner. Exams disadvantage women in introductory biology. *PLoS One*, 12(10):e0186419, 2017.

[17] Mahzarin R Banaji and Anthony G Greenwald. *Blindspot: Hidden biases of good people*. Bantam, 2016.

[18] A. Bandura. Social cognitive theory of self-regulation. *Organ. Behav. Hum. Decis. Process*, 50:248–287, 1991.

[19] A. Bandura. *Self-Efficacy: The Exercise of Control*. Freeman, New York, 1997.

[20] Albert Bandura. Self-efficacy mechanism in human agency. *Am. Psychol*, 37(2):122, 1982.

[21] Lei Bao, Tianfan Cai, Kathy Koenig, Kai Fang, Jing Han, Jing Wang, Qing Liu, Lin Ding, Lili Cui, Ying Luo, et al. Learning and scientific reasoning. *Science*, 323(5914):586–587, 2009.

[22] Lei Bao and Kathleen Koenig. Physics education research for 21st century learning. *Disciplinary and Interdisciplinary Science Education Research*, 1(2):2, 2019.

[23] Ramón Barthelemy, Melinda McCormick, and Charles Henderson. Barriers beyond equity: an exploratory study of women graduate students' career pathways in astronomy. *International Journal of Gender, Science and Technology*, 7(1):57–73, 2015.

[24] Ramón S Barthelemy, Melinda McCormick, and Charles Henderson. Gender discrimination in physics and astronomy: Graduate student experiences of sexism and gender microaggressions. *Phys. Rev. Phys. Educ. Res.*, 12(2):020119, 2016.

[25] R Beichner. History and evolution of active learning spaces. In *Active Learning Spaces*, pages 9–16. Jossey-Bass, 2014.

[26] Robert J Beichner, Jeffery M Saul, David S Abbott, Jeanne J Morse, Duane Deardorff, Rhett J Allain, Scott W Bonham, Melissa H Dancy, and John S Risley. The student-centered activities for large enrollment undergraduate programs (SCALE-UP) project. *Research-based reform of university physics*, 1(1):2–39, 2007.

[27] John W Belcher. Studio physics at MIT. *Interface*, pages 58–64, 2001.

[28] S D Bergin, C Murphy, and A Ni Shuilleabhain. Exploring problem-based cooperative learning in undergraduate physics labs: student perspectives. *Eur. J. Phys.*, 39(2):025703, Jan 2018.

[29] Kevin R. Binning, Nancy Kaufmann, Erica M. McGreevy, Omid Fotuhi, Susie Chen, Emily Marshman, Z. Yasemin Kalender, Lisa Limeri, Laura Betancur, and Chandralekha Singh. Changing social contexts to foster equity in college science courses: An ecological-belonging intervention. *Psychological Science*, 31(9):1059–1070, 2020.

[30] Robin Bjorkquist, Abigail M. Bogdan, Nicole L. Campbell, Mary Chessey, Geraldine L. Cochran, Beth Cunningham, Jessica N. Esquivel, Laura Gladstone, Natalie M. Gosnell, Sathya Guruswamy, Kelsey M. Hallinen, Candace Harris, Angela Johnson, Jolene L. Johnson, Christine Jones, Regina A. Jorgenson, Laura McCullough, Marta D. McNeese, Tennille D. Presley, Nicole Quist, Arlisa Richardson, Sally Seidel,

and Chandralekha Singh. Women in physics in the United States: Reaching toward equity and inclusion. *AIP Conference Proceedings*, 2109(1):050040, 2019.

[31] A. A. Bless. Cook-book laboratory work. *The American Physics Teacher*, 1:88–89, 1933.

[32] Erik Bodegom, Erik Jensen, and David Sokoloff. Adapting realtime physics for distance learning with the iolab. *The Physics Teacher*, 57(6):382–386, 2019.

[33] F. Bouquet, J. Bobroff, M. Fuchs-Gallezot, and L. Maurines. Project-based physics labs using low-cost open-source hardware. *American Journal of Physics*, 85(3):216–222, 2017.

[34] D. Branning, S. Bhandari, and M. Beck. Low-cost coincidence-counting electronics for undergraduate quantum optics. *American Journal of Physics*, 77(7):667–670, 2009.

[35] Eric Brewe. Modeling theory applied: Modeling instruction in introductory physics. *Am. J. Phys.*, 76(12):1155–1160, 2008.

[36] David T. Brookes, Yuehai Yang, and Binod Nainabasti. Social positioning in small group interactions in an investigative science learning environment physics class. *Phys. Rev. Phys. Educ. Res.*, 17:010103, Jan 2021.

[37] Amy Bug. Has feminism changed physics? *Signs*, 28(3):881–899, 2003.

[38] M. Burawoy. *The Extended Case Method: Four Countries, Four Decades, Four Great Transformations, and One Theoretical Tradition*. University of California Press, 2009.

[39] Marie Buscatto. Practising reflexivity in ethnography. In David Silverman, editor, *Qualitative Research*, chapter 9, pages 137–151. Sage, 1997.

[40] Kristine E. Callan, Bethany Wilcox, and Wendy Adams. Testing group composition within a studio learning environment. In *Physics Education Research Conference 2017*, PER Conference, pages 72–75, Cincinnati, OH, July 26-27 2017.

[41] Michelle Madsen Camacho and Susan M Lord. Microaggressions in engineering education: Climate for asian, latina and white women. In *2011 Frontiers in Education Conference (FIE)*. IEEE, 2011.

[42] Enrico Gianfranco Campari, Manuel Barbetta, Sylvie Braibant, Nitya Cuzzuol, Alessandro Gesuato, Leonardo Maggiore, Federico Marulli, Giovanni Venturoli, and Cristian Vignali. Physics laboratory at home during the COVID-19 pandemic. *The Physics Teacher*, 59(1):68–71, 2021.

[43] Emanuela Carleschi, Anna Chrysostomou, Alan S Cornell, and Wade Naylor. Does transitioning to online classes mid-semester affect conceptual understanding? *arXiv preprint arXiv:2101.09908*, 2021.

[44] Heidi B Carlone and Angela Johnson. Understanding the science experiences of successful women of color: Science identity as an analytic lens. *J. Res. Sci. Teach.*, 44(8):1187–1218, 2007.

[45] Timothy Chambers. *Three pedagogical approaches to introductory physics labs and their effects on student learning outcomes*. PhD thesis, University of Arizona, 2014.

[46] Elaine Chapman. Alternative approaches to assessing student engagement rates. *Practical Assessment, Research, and Evaluation*, 8(1):13, 2002.

[47] Michelene TH Chi, Paul J Feltovich, and Robert Glaser. Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5(2):121–152, 1981.

[48] Jacquelyn J. Chini and Ahlam Al-Rawi. Alignment of TAs' beliefs with practice and student perception. *AIP Conference Proceedings*, 1513(1):98–101, 2013.

[49] R. E. Clark, D. F. Feldon, J. G. van Merriënboer, K. Yates, and S. Early. Handbook of research on educational communications and technology. In J. M. Spector, M. D. Merrill, J. G. van Merriënboer, and M. P. Driscoll, editors, *Cognitive Task Analysis*. Lawrence Erlbaum Associates, 2007.

[50] Russell Clark. *Introduction to Laboratory Physics*. Kendall Hunt, 3 edition, 2012.

[51] Eleanor W. Close, Jessica Conn, and Hunter G. Close. Becoming physics people: Development of integrated physics identity through the learning assistant experience. *Phys. Rev. Phys. Educ. Res.*, 12:010109, Feb 2016.

[52] Geraldine L Cochran and Mel S Sabella. Understanding and encouraging effective collaboration in introductory physics courses. In *AIP Conference Proceedings*, volume 1064, pages 95–98. AIP, 2008.

[53] Elisabeth G Cohen and Rachel A Lotan. *Designing Groupwork: Strategies for the Heterogeneous Classroom*. Teachers College Press, 3 edition, 2014.

[54] Jacob Cohen. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1):37–46, 1960.

[55] Jacob Cohen. *Statistical Power Analysis for the Behavioral Sciences*. Academic Press, 2013.

[56] Allan Collins, John Seely Brown, and Ann Holum. Cognitive apprenticeship: Making thinking visible. *American Educator*, 15(3):6–11, 1991.

[57] Sehoya Cotner and Cissy J Ballen. Can mixed assessment methods make biology classes more equitable? *PLoS One*, 12(12):e0189610, 2017.

[58] John W Creswell and J David Creswell. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. Sage Publications, Thousand Oaks, CA, 2017.

[59] Kevin Crowley, Maureen A Callanan, Harriet R Tenenbaum, and Elizabeth Allen. Parents explain more often to boys than to girls during shared scientific thinking. *Psychol. Sci.*, 12(3):258–261, 2001.

[60] Daniel A. Dale, Jessica Sutter, and Dylan Kloster. Asking real-world questions with inquiry-based labs. *The Physics Teacher*, 57(8):547–550, 2019.

[61] Anna Teresia Danielsson and Cedric Linder. Learning in physics by doing laboratory work: Towards a new conceptual framework. *Gend. Educ.*, 21(2):129–144, 2009.

[62] James Day, Jared B. Stang, N. G. Holmes, Dhaneesh Kumar, and D. A. Bonn. Gender gaps and gendered action in a first-year physics laboratory. *Phys. Rev. Phys. Educ. Res.*, 12:020104, Aug 2016.

[63] George DeBeck, Sam Settelmeyer, Sissi Li, and Dedra Demaree. TA beliefs in a SCALE-UP style classroom. *AIP Conference Proceedings*, 1289(1):121–124, 2010.

[64] G.E. DeBoer. *A History of Ideas in Science Education: Implications for Practice*. Teachers College Press, 1991.

[65] Seth DeVore, Alexandre Gauthier, Jeremy Levy, and Chandralekha Singh. Development and evaluation of a tutorial to improve students' understanding of a lock-in amplifier. *Phys. Rev. Phys. Educ. Res.*, 12:020127, Sep 2016.

[66] Seth DeVore, Alexandre Gauthier, Jeremy Levy, and Chandralekha Singh. Improving student understanding of lock-in amplifiers. *American Journal of Physics*, 84(1):52–56, 2016.

[67] Seth DeVore, Emily Marshman, and Chandralekha Singh. Challenge of engaging all students via self-paced interactive electronic learning tutorials for introductory physics. *Phys. Rev. Phys. Educ. Res.*, 13:010127, May 2017.

[68] Matthew Dew, Lewis Ford, Dawson T Nodurft, Tatiana Erukhimova, and Jonathan Perry. Student responses to changes in introductory physics learning due to the covid-19 pandemic. *The Physics Teacher*, 59(3):162–165, 2021.

[69] John Dewey. *Experiential Education.* Collier, New York, 1938.

[70] Yehudit Judy Dori and John Belcher. How does technology-enabled active learning affect undergraduate students' understanding of electromagnetism concepts? *The Journal of the Learning Sciences*, 14(2):243–279, 2005.

[71] Danny Doucette. Lab TA TRAINING, 2019. http://labtatraining.com.

[72] Danny Doucette, Russell Clark, and Chandralekha Singh. What's happening in traditional and inquiry-based introductory labs? An integrative analysis at a large research university. In *Physics Education Research Conference 2018*, Washington, DC, August 1-2 2018.

[73] Danny Doucette, Russell Clark, and Chandralekha Singh. All aboard! Challenges and successes in training lab TAs. In *Physics Education Research Conference 2019*, 2020.

[74] Danny Doucette, Russell Clark, and Chandralekha Singh. Hermione and the secretary: How gendered task division in introductory physics labs can disrupt equitable learning. *European Journal of Physics*, 41(3):035702, apr 2020.

[75] Danny Doucette, Russell Clark, and Chandralekha Singh. Professional development combining cognitive apprenticeship and expectancy-value theories improves lab

teaching assistants' instructional views and practices. *Phys. Rev. Phys. Educ. Res.*, 16:020102, Jul 2020.

[76] Danny Doucette, Russell Clark, and Chandralekha Singh. What makes a good physics lab partner? In *Physics Education Research Conference 2020*, PER Conference, pages 124–130, Virtual Conference, July 22-23 2020.

[77] Danny Doucette, Russell Clark, and Chandralekha Singh. What makes a good physics lab partner? In *Physics Education Research Conference 2020*, PER Conference, pages 124–130, Virtual Conference, July 22-23 2020.

[78] Danny Doucette, Brian D'Urso, and Chandralekha Singh. Lessons from transforming second-year honors physics lab. *American Journal of Physics*, 88(10):838–844, 2020.

[79] Danny Doucette and Chandralekha Singh. Why are there so few women in physics? Reflections on the experiences of two women. *The Physics Teacher*, 58(5):297–300, 2020.

[80] Danny Doucette and Chandralekha Singh. Expansive framing produces more vivid introductory physics labs. In Maria Teresa Caccamo, editor, *Physics Education for Students: An Interdisciplinary Approach*. Bentham Science, Sharjah, 2021. in press.

[81] Danny Doucette and Chandralekha Singh. Views of female students who played the role of group leaders in introductory physics labs. *Eur. J. Phys.*, 2021. https://doi.org/10.1088/1361-6404/abd597.

[82] Scott S. Douglas, John M. Aiken, Shih-Yin Lin, Edwin F. Greco, Emily Alicea-Muñoz, and Michael F. Schatz. Peer assessment of student-produced mechanics lab report videos. *Phys. Rev. Phys. Educ. Res.*, 13:020126, Nov 2017.

[83] Dimitri R. Dounas-Frazer and H. J. Lewandowski. Electronics lab instructors' approaches to troubleshooting instruction. *Phys. Rev. Phys. Educ. Res.*, 13:010102, Jan 2017.

[84] Dimitri R. Dounas-Frazer and Daniel L. Reinholz. Attending to lifelong learning skills through guided reflection in a physics class. *American Journal of Physics*, 83(10):881–891, 2015.

[85] Dimitri R. Dounas-Frazer, Jacob T. Stanley, and H. J. Lewandowski. Student ownership of projects in an upper-division optics laboratory course: A multiple case study of successful experiences. *Phys. Rev. Phys. Educ. Res.*, 13:020136, Dec 2017.

[86] Erin M. Duffy and Melanie M. Cooper. Assessing TA buy-in to expectations and alignment of actual teaching practices in a transformed general chemistry laboratory course. *Chem. Educ. Res. Pract.*, 21:189–208, 2020.

[87] R Dufresne, J Mestre, T Thaden-Koch, W Gerace, and W Leonard. When transfer fails: Effect of knowledge, expectations and observations on transfer in physics. In *Transfer of learning: Research and perspectives*, pages 155–215. 2005.

[88] K Dunnett, M N Gorman, and P A Bartlett. Assessing first-year undergraduate physics students' laboratory practices: seeking to encourage research behaviours. *Eur. J. Phys.*, 40(1):015702, dec 2018.

[89] Brian D'Urso. Pythics, 2015. https://github.com/dursobr/Pythics.

[90] Ben Van Dusen, Laurie Langdon, and Valerie Otero. Learning assistant supported student outcomes (LASSO) study initial findings. In *Physics Education Research Conference 2015*, PER Conference, pages 343–346, College Park, MD, July 29-30 2015.

[91] Alice H Eagly. When passionate advocates meet research on diversity, does the honest broker stand a chance? *Journal of Social Issues*, 72(1):199–222, 2016.

[92] Jacquelynne S. Eccles and Allan Wigfield. Motivational beliefs, values, and goals. *Annual Review of Psychology*, 53(1):109–132, 2002.

[93] R. A. Engle. Framing interactions to foster generative learning: A situative explanation of transfer in a community of learners classroom. *Journal of the Learning Sciences*, 15:451–498, 2006.

[94] Randi A Engle, Diane P Lam, Xenia S Meyer, and Sarah E Nix. How does expansive framing promote transfer? Several proposed explanations and a research agenda for investigating them. *Educational Psychologist*, 47(3):215–231, 2012.

[95] Randi A Engle, Phi D Nguyen, and Adam Mendelson. The influence of framing on transfer: Initial evidence from a tutoring experiment. *Instructional Science*, 39(5):603–628, 2011.

[96] E. Etkina and A. Van Heuvelen. Investigative science learning environment - a science process approach to learning physics. In *Research-Based Reform of University Physics*, volume 1. April 2007.

[97] Joseph L Fleiss, Bruce Levin, and Myunghee Cho Paik. *Statistical methods for rates and proportions*. John Wiley & Sons, 2013.

[98] Michael FJ Fox, Alexandra Werth, Jessica R Hoehn, and HJ Lewandowski. Teaching labs during a pandemic: Lessons from spring 2020 and an outlook for the future. *arXiv preprint arXiv:2007.01271*, 2020.

[99] Scott Freeman, Sarah L. Eddy, Miles McDonough, Michelle K. Smith, Nnadozie Okoroafor, Hannah Jordt, and Mary Pat Wenderoth. Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, 111(23):8410–8415, 2014.

[100] Mary P Freier. The librarian in Rowling's Harry Potter series. *CLCWeb: Comparative Literature and Culture*, 16(3):6, 2014.

[101] D Galan, R Heradio, L de la Torre, S Dormido, and F Esquembre. The experiment editor: supporting inquiry-based learning with virtual labs. *Eur. J. Phys.*, 38(3):035702, mar 2017.

[102] Calin Galeriu, Cheryl Letson, and Geoffrey Esper. An Arduino investigation of the RC circuit. *The Physics Teacher*, 53(5):285–288, 2015.

[103] E. J. Galvez, Charles H. Holbrow, M. J. Pysher, J. W. Martin, N. Courtemanche, L. Heilig, and J. Spencer. Interference with correlated photons: Five quantum mechanics experiments for undergraduates. *American Journal of Physics*, 73(2):127–140, 2005.

[104] Enrique Galvez and Chandralekha Singh. Introduction to the theme issue on experiments and laboratories in physics education. *American Journal of Physics*, 78(5):453–454, 2010.

[105] Andrew Gavrin. Physics students' reactions to an abrupt shift in instruction during the covid-19 pandemic. In *Physics Education Research Conference 2020*, PER Conference, pages 167–172, Virtual Conference, July 22-23 2020.

[106] James Paul Gee. Identity as an analytic lens for research in education. *Rev. Res. Educ.*, 25:99–125, 2000.

[107] Benjamin D. Geller, Chandra Turpen, and Catherine H. Crouch. Sources of student engagement in Introductory Physics for Life Sciences. *Phys. Rev. Phys. Educ. Res.*, 14:010118, Apr 2018.

[108] Andrew Gelman and Jennifer Hill. *Data Analysis using Regression and Multilevel/Hierarchical Models*. Cambridge University Press, 2006.

[109] Elizabeth A George, Maan Jiang Broadstock, and Jesús Vázquez Abad. Learning energy, momentum, and conservation concepts with computer support in an undergraduate physics laboratory. In *Fourth International Conference of the Learning Sciences*, pages 2–3, 2013.

[110] John A Gilreath and Timothy F Slater. Training graduate teaching assistants to be better undergraduate physics educators. *Phys. Educ*, 29(4):200–203, jul 1994.

[111] Allison Godwin, Geoff Potvin, Zahra Hazari, and Robynne Lock. Identity, critical agency, and engineering: An affective model for predicting engineering as a career choice. *J. Eng. Educ.*, 105(2):312–340, 2016.

[112] Renee Michelle Goertzen, Rachel E. Scherr, and Andrew Elby. Accounting for tutorial teaching assistants' buy-in to reform instruction. *Phys. Rev. ST Phys. Educ. Res.*, 5:020109, Dec 2009.

[113] Renee Michelle Goertzen, Rachel E. Scherr, and Andrew Elby. Respecting tutorial instructors' beliefs and experiences: A case study of a physics teaching assistant. *Phys. Rev. ST Phys. Educ. Res.*, 6:020125, Dec 2010.

[114] Renee Michelle Goertzen, Rachel E. Scherr, and Andrew Elby. Tutorial teaching assistants in the classroom: Similar teaching behaviors are supported by varied beliefs about teaching and learning. *Phys. Rev. ST Phys. Educ. Res.*, 6:010105, Apr 2010.

[115] Allison J Gonsalves, Anna Danielsson, and Helena Pettersson. Masculinities and experimental practices in physics: The view from three case studies. *Phys. Rev. Phys. Educ. Res.*, 12(2):020120, 2016.

[116] Melanie Good, Alexandru Maries, and Chandralekha Singh. Impact of traditional or evidence-based active-engagement instruction on introductory female and male

students' attitudes and approaches to physics problem solving. *Phys. Rev. Phys. Educ. Res.*, 15:020129, Sep 2019.

[117] Melanie Good, Emily Marshman, Edit Yerushalmi, and Chandralekha Singh. Physics teaching assistants' views of different types of introductory problems: Challenge of perceiving the instructional benefits of context-rich and multiple-choice problems. *Phys. Rev. Phys. Educ. Res.*, 14:020120, Nov 2018.

[118] Melanie Good, Emily Marshman, Edit Yerushalmi, and Chandralekha Singh. Graduate teaching assistants' views of broken-into-parts physics problems: Preference for guidance overshadows development of self-reliance in problem solving. *Phys. Rev. Phys. Educ. Res.*, 16:010128, May 2020.

[119] Maithreyi Gopalakrishnan and Markus Gühr. A low-cost mirror mount control system for optics setups. *American Journal of Physics*, 83(2):186–190, 2015.

[120] Chris Gosling. Identity as a research lens in science and physics education. *Journal of Belonging, Identity, Language, and Diversity*, 1(1):62–74, 2018.

[121] Anneke L. Gretton, Terry Bridges, and James M. Fraser. Transforming physics educator identities: TAs help TAs become teaching professionals. *Am. J. Phys.*, 85(5):381–391, 2017.

[122] Sarah S Grover, Tiffany A Ito, and Bernadette Park. The effects of gender composition on women's experience in math work groups. *Journal of Personality and Social Psychology*, 112(6):877, 2017.

[123] Sandra Harding. Strong objectivity: A response to the new objectivity question. *Synthese*, 104(3):331–349, 1995.

[124] Zahra Hazari, Gerhard Sonnert, Philip M Sadler, and Marie-Claire Shanahan. Connecting high school physics experiences, outcome expectations, physics identity, and physics career choice: A gender study. *J. Res. Sci. Teach.*, 47(8):978–1003, 2010.

[125] Patricia Heller and Kenneth Heller. *Cooperative Group Problem Solving in Physics.* Brooks/Cole Publishing Company, 2001.

[126] Patricia Heller and Mark Hollabaugh. Teaching problem solving through cooperative grouping. Part 2: Designing problems and structuring groups. *Am. J. Phys.*, 60(7):637–644, 1992.

[127] Eden Hennessey, Joanne Cole, Prajval Shastri, Jessica Esquivel, Chandralekha Singh, Rosie Johnson, and Shohini Ghose. Workshop report: intersecting identities: gender and intersectionality in physics. *AIP Conference Proceedings*, 2109(1):040001, 2019.

[128] Xochith Herrera, Jayson Nissen, and Ben Van Dusen. Student outcomes across collaborative-learning environments. In *Physics Education Research Conference 2018*, PER Conference, Washington, DC, August 1-2 2018.

[129] David Hestenes and Malcolm Wells. A Mechanics Baseline Test. *The Physics Teacher*, 30(3):159–166, 1992.

[130] David Hestenes, Malcolm Wells, and Gregg Swackhamer. Force concept inventory. *Phys. Teach.*, 30(3):141–158, March 1992.

[131] Eric Hickok. Teaching practices of graduate teaching assistants. Master's thesis, San José State University, 2016.

[132] Suzanne Hidi and K Ann Renninger. The four-phase model of interest development. *Educational psychologist*, 41(2):111–127, 2006.

[133] Robert C. Hilborn. Physics and the revised Medical College Admission Test. *American Journal of Physics*, 82(5):428–433, 2014.

[134] N. Holmes, J. Olsen, J. Thomas, and C. Wieman. Value added or misattributed? A multi-institution study on the educational benefit of labs for reinforcing physics content. *Phys. Rev. PER*, 13:010129, May 2017.

[135] N. Holmes and C. Wieman. Introductory physics labs: We can do better. *Physics Today*, 71(1):38–45, 2018.

[136] N. G. Holmes and H. J. Lewandowski. Investigating the landscape of physics laboratory instruction across north america. *Phys. Rev. Phys. Educ. Res.*, 16:020162, Dec 2020.

[137] N. G. Holmes and Emily M. Smith. Operationalizing the AAPT learning goals for the lab. *The Physics Teacher*, 57(5):296–299, 2019.

[138] N. G. Holmes, Carl E. Wieman, and D. A. Bonn. Teaching critical thinking. *Proceedings of the National Academy of Sciences*, 112(36):11199–11204, 2015.

[139] Natasha G. Holmes, Matthew "Sandy" Martinuk, Joss Ives, and Mya Warren. Teaching assistant professional development by and for TAs. *The Physics Teacher*, 51(4):218–219, 2013.

[140] Natasha Grace Holmes. *Structured quantitative inquiry labs: Developing critical thinking in the introductory physics laboratory.* PhD thesis, University of British Columbia, 2014.

[141] N.G. Holmes, Ido Roll, and D.A. Bonn. Participating in the physics lab: Does gender matter? *Physics in Canada, Special Issue on Physics Education Research*, 70(2):84–86, 2014.

[142] Jen-Feng Hsu, Shonali Dhingra, and Brian D'Urso. Design and construction of a cost-efficient Arduino-based mirror galvanometer system for scanning optical microscopy. *American Journal of Physics*, 85(1):68–75, 2017.

[143] Dehui Hu and Benjamin M. Zwickl. Examining students' views about validity of experiments: From introductory to Ph.D. students. *Phys. Rev. Phys. Educ. Res.*, 14:010121, Apr 2018.

[144] Dehui Hu, Benjamin M. Zwickl, Bethany R. Wilcox, and H. J. Lewandowski. Qualitative investigation of students' views about experimental physics. *Phys. Rev. Phys. Educ. Res.*, 13:020134, Nov 2017.

[145] Chris S Hulleman, Olga Godes, Bryan L Hendricks, and Judith M Harackiewicz. Enhancing interest and performance with a utility value intervention. *Journal of Educational Psychology*, 102(4):880, 2010.

[146] Anne Barrie Hunter, Sandra L. Laursen, and Elaine Seymour. Becoming a scientist: The role of undergraduate research in students' cognitive, personal, and professional development. *Science Education*, 91(1):36–74, 1 2007.

[147] J Huntula, M D Sharma, I Johnston, and R Chitaree. A framework for laboratory pre-work based on the concepts, tools and techniques questioning method. *European Journal of Physics*, 32(5):1419–1430, Aug 2011.

[148] Simone Hyater-Adams, Claudia Fracchiolla, Noah Finkelstein, and Kathleen Hinko. Critical look at physics identity: An operationalized framework for examining race and physics identity. *Phys. Rev. Phys. Educ. Res.*, 14(1):010132, 2018.

[149] Paul W. Irving and Eleanor C. Sayre. Conditions for building a community of practice in an advanced physics laboratory. *Phys. Rev. ST Phys. Educ. Res.*, 10:010109, Mar 2014.

[150] Hyewon Jang, Nathaniel Lasry, Kelly Miller, and Eric Mazur. Collaborative exams: Cheating? Or learning? *American Journal of Physics*, 85(3):223–227, 2017.

[151] Angela C. Johnson. Unintended consequences: How science professors discourage women of color. *Science Education*, 91(5):805–821, 2007.

[152] David W Johnson, Roger T Johnson, and Karl A Smith. Cooperative learning: Improving university instruction by basing practice on validated theory. *Journal on Excellence in University Teaching*, 25(4):1–26, 2014.

[153] Molly Johnson. Facilitating high quality student practice in introductory physics. *American Journal of Physics*, 69(S1):S2–S11, 2001.

[154] E. Leonard Jossem. Resource letter EPGA-1: The education of physics graduate assistants. *Am. J. Phys.*, 68(6):502–512, 2000.

[155] Z. Yasemin Kalender, Emily Marshman, Christian D. Schunn, Timothy J. Nokes-Malach, and Chandralekha Singh. Gendered patterns in the construction of physics identity from motivational factors. *Phys. Rev. Phys. Educ. Res.*, 15:020119, Aug 2019.

[156] Z. Yasemin Kalender, Emily Marshman, Christian D. Schunn, Timothy J. Nokes-Malach, and Chandralekha Singh. Why female science, technology, engineering, and mathematics majors do not identify with physics: They do not think others see them that way. *Phys. Rev. Phys. Educ. Res.*, 15:020148, Dec 2019.

[157] Z. Yasemin Kalender, Emily Marshman, Christian D. Schunn, Timothy J. Nokes-Malach, and Chandralekha Singh. Damage caused by women's lower self-efficacy on physics learning. *Phys. Rev. Phys. Educ. Res.*, 16:010118, Apr 2020.

[158] Calvin S. Kalman, Marina Milner-Bolotin, and Tetyana Antimirova. Comparison of the effectiveness of collaborative groups and peer instruction in a large introductory physics course for science majors. *Canadian Journal of Physics*, 88(5):325–332, 2010.

[159] Calvin S. Kalman, Mandana Sobhanzadeh, Robert Thompson, Ahmed Ibrahim, and Xihui Wang. Combination of interventions can change students' epistemological beliefs. *Phys. Rev. ST Phys. Educ. Res.*, 11:020136, Dec 2015.

[160] Manu Kapur. Productive failure. *Cognition and Instruction*, 26(3):379–424, 2008.

[161] Nafis I Karim, Alexandru Maries, and Chandralekha Singh. Do evidence-based active-engagement courses reduce the gender gap in introductory physics? *Eur. J. Phys.*, 39(2):025701, 2018.

[162] Nafis I. Karim, Alexandru Maries, and Chandralekha Singh. Exploring one aspect of pedagogical content knowledge of teaching assistants using the Conceptual Survey of Electricity and Magnetism. *Phys. Rev. Phys. Educ. Res.*, 14:010117, Apr 2018.

[163] Nafis I Karim, Alexandru Maries, and Chandralekha Singh. Impact of evidence-based flipped or active-engagement non-flipped courses on student performance in introductory physics. *Canadian Journal of Physics*, 96(4):411–419, 2018.

[164] Alison King. From sage on the stage to guide on the side. *College Teaching*, 41(1):30–35, 1993.

[165] L Kirkup, S Johnson, E Hazel, R W Cheary, D C Green, P Swift, and W Holliday. Designing a new physics laboratory programme for first-year engineering students. *Phys. Educ*, 33(4):258–265, Jul 1998.

[166] Les Kirkup, Jenny Pizzica, Katrina Waite, and Lakshmi Srinivasan. Realizing a framework for enhancing the laboratory experiences of non-physics majors: from pilot to large-scale implementation. *European Journal of Physics*, 31(5):1061–1070, Jul 2010.

[167] Barry Kirwan and Les K Ainsworth. *A Guide to Task Analysis: The Task Analysis Working Group*. CRC Press, 1992.

[168] Pascal Klein, Lana Ivanjek, Merten Nikolay Dahlkemper, Katarina Jeličić, M-A Geyer, Stefan Küchemann, and Ana Susac. Studying physics during the COVID-19 pandemic: Student assessments of learning achievement, perceived effectiveness of online recitations, and online laboratories. *Physical Review Physics Education Research*, 17(1):010117, 2021.

[169] Kathleen Koenig, Krista E Wood, Larry J Bortner, and Lei Bao. Modifying traditional labs to target scientific reasoning. *J. of Coll. Sci. Teach.*, 48(5):28–35, 2019.

[170] Kathleen M. Koenig, Robert J. Endorf, and Gregory A. Braun. Effectiveness of different tutorial recitation teaching methods and its implications for TA training. *Phys. Rev. ST Phys. Educ. Res.*, 3:010104, May 2007.

[171] Ismo T Koponen and Terhi Mäntylä. Generative role of experiments in physics and in teaching physics: A suggestion for epistemological reconstruction. *Science & Education*, 15(1):31–54, 2006.

[172] Gerd Kortemeyer, Edwin Kashy, Walter Benenson, and Wolfgang Bauer. Experiences using the open-source learning content management and assessment system LON-CAPA in introductory physics courses. *American Journal of Physics*, 76(4):438–444, 2008.

[173] Gerd Kortemeyer and Anna Fritchie Kortemeyer. The nature of collaborations on programming assignments in introductory physics courses: a case study. *Eur. J. Phys.*, 39(5):055705, aug 2018.

[174] Melinda L Kreth. A survey of the co-op writing experiences of recent engineering graduates. *IEEE Trans. Prof. Commun.*, 43(2):137–152, 2000.

[175] J. Richard Landis and Gary G. Koch. The measurement of observer agreement for categorical data. *Biometrics*, 33(1):159–174, 1977.

[176] Nathaniel Lasry, Elizabeth Charles, and Chris Whittaker. When teacher-centered instructors are assigned to student-centered classrooms. *Phys. Rev. ST Phys. Educ. Res.*, 10:010116, May 2014.

[177] C. M. Lavelle. Gamma ray spectroscopy with Arduino UNO. *American Journal of Physics*, 86(5):384–394, 2018.

[178] F. Lawrenz, P. Heller, R. Keith, and K. Heller. Training the teaching assistant. *Journal of College Science Teaching*, 22(2):106–9, 1992.

[179] Priscilla Laws. Workshop physics: Learning introductory physics by doing it. *Change: The Magazine of Higher Learning*, 23(4):20–27, 1991.

[180] Priscilla W Laws, Pamela J Rosborough, and Frances J Poodry. Women's responses to an activity-based introductory physics program. *Am. J. Phys.*, 67(S1):32–37, 1999.

[181] Louis Leblond and Melissa Hicks. Designing laboratories for online instruction using the iOLab device. *The Physics Teacher*, 59(5):351–355, 2021.

[182] Norman G Lederman et al. Nature of science: Past, present, and future. *Handbook of Research on Science Education*, 2:831–879, 2007.

[183] Olivia Levrini, Anna De Ambrosis, Sabine Hemmer, Antti Laherto, Massimiliano Malgieri, Ornella Pantano, and Giulia Tasquier. Understanding first-year students' curiosity and interest about physics—lessons learned from the HOPE project. *European Journal of Physics*, 38(2):025701, dec 2016.

[184] Shih-Yin Lin, Charles Henderson, William Mamudi, Chandralekha Singh, and Edit Yerushalmi. Teaching assistants' beliefs regarding example solutions in introductory physics. *Phys. Rev. ST Phys. Educ. Res.*, 9(1):010120, 2013.

[185] Shih-Yin Lin and Chandralekha Singh. Challenges in using analogies. *The Physics Teacher*, 49(8):512–513, 2011.

[186] Shih-Yin Lin and Chandralekha Singh. Using isomorphic problems to learn introductory physics. *Phys. Rev. ST Phys. Educ. Res.*, 7:020104, Aug 2011.

[187] Shih-Yin Lin and Chandralekha Singh. Using an isomorphic problem pair to learn introductory physics: Transferring from a two-step problem to a three-step problem. *Phys. Rev. ST Phys. Educ. Res.*, 9:020114, Oct 2013.

[188] Shih-Yin Lin and Chandralekha Singh. Effect of scaffolding on helping introductory physics students solve quantitative problems involving strong alternative conceptions. *Phys. Rev. ST Phys. Educ. Res.*, 11:020105, Aug 2015.

[189] Marcia C. Linn, Erin Palmer, Anne Baranger, Elizabeth Gerard, and Elisa Stone. Undergraduate research experiences: Impacts and opportunities. *Science*, 347(6222):627, 2015.

[190] Robynne M. Lock, Zahra Hazari, and Geoff Potvin. Impact of out-of-class science and engineering activities on physics identity and career intentions. *Phys. Rev. Phys. Educ. Res.*, 15:020137, Oct 2019.

[191] Mercedes Lorenzo, Catherine H. Crouch, and Eric Mazur. Reducing the gender gap in the physics classroom. *American Journal of Physics*, 74(2):118–122, 2006.

[192] Julie A Luft, Josepha P Kurdziel, Gillian H Roehrig, and Jessica Turner. Growing a garden without water: Graduate teaching assistants in introductory science laboratories at a doctoral/research university. *Journal of Research in Science Teaching*, 41(3):211–233, 2004.

[193] Mark J Macgowan. A measure of engagement for social group work: The groupwork engagement measure (GEM). *Journal of Social Service Research*, 23(2):17–37, 1997.

[194] D Soyini Madison. *Critical Ethnography: Method, Ethics, and Performance*. Sage, 2011.

[195] Gwen C. Marchand and Gita Taasoobshirazi. Stereotype threat and women's performance in physics. *International Journal of Science Education*, 35(18):3050–3061, 2013.

[196] Alexandru Maries. Preparing the next generation of educators. In J. J. Mintzes and E. Walter, editor, *Active Learning in College Science*. Springer, 2020.

[197] Alexandru Maries, Nafis I Karim, and Chandralekha Singh. Is agreeing with a gender stereotype correlated with the performance of female students in introductory physics? *Phys. Rev. Phys. Educ. Res.*, 14(2):020119, 2018.

[198] Alexandru Maries, Shih-Yin Lin, and Chandralekha Singh. Challenges in designing appropriate scaffolding to improve students' representational consistency: The case of a Gauss's law problem. *Phys. Rev. Phys. Educ. Res.*, 13:020103, Aug 2017.

[199] Alexandru Maries and Chandralekha Singh. Exploring one aspect of pedagogical content knowledge of teaching assistants using the Test of Understanding Graphs in Kinematics. *Phys. Rev. ST Phys. Educ. Res.*, 9:020120, Nov 2013.

[200] Alexandru Maries and Chandralekha Singh. Teaching assistants' performance at identifying common introductory student difficulties in mechanics revealed by the Force Concept Inventory. *Phys. Rev. Phys. Educ. Res.*, 12:010131, May 2016.

[201] E. Marshman, Z. Kalender, C. Schunn, T. Nokes-Malach, and C. Singh. A longitudinal analysis of students' motivational characteristics in introductory physics courses: Gender differences. *Canadian Journal of Physics*, 96(4):391–405, 2018.

[202] Emily Marshman, Alexandru Maries, Ryan T Sayer, Charles Henderson, Edit Yerushalmi, and Chandralekha Singh. Physics postgraduate teaching assistants'

grading approaches: Conflicting goals and practices. *European Journal of Physics*, 41(5):055701, Aug 2020.

[203] Emily Marshman, Ryan Sayer, Charles Henderson, and Chandralekha Singh. Contrasting grading approaches in introductory physics and quantum mechanics: The case of graduate teaching assistants. *Phys. Rev. Phys. Educ. Res.*, 13:010120, May 2017.

[204] Emily Marshman, Ryan Sayer, Charles Henderson, Edit Yerushalmi, and Chandralekha Singh. The challenges of changing teaching assistants' grading practices: Requiring students to show evidence of understanding. *Canadian Journal of Physics*, 96(4):420–437, 2018.

[205] Emily M Marshman, Z Yasemin Kalender, Timothy Nokes-Malach, Christian Schunn, and Chandralekha Singh. Female students with A's have similar physics self-efficacy as male students with C's in introductory courses: A cause for alarm? *Phys. Rev. Phys. Educ. Res.*, 14(2):020123, 2018.

[206] Andrew Mason and Chandralekha Singh. Helping students learn effective problem solving strategies by reflecting with peers. *American Journal of Physics*, 78(7):748–754, 2010.

[207] Andrew Mason and Chandralekha Singh. Assessing expertise in introductory physics using categorization task. *Phys. Rev. ST Phys. Educ. Res.*, 7:020110, Oct 2011.

[208] Andrew Mason and Chandralekha Singh. Using categorization of problems as an instructional tool to help introductory students learn physics. *Physics Education*, 51(2):025009, 2016.

[209] Andrew Mason, Edit Yerushalmi, Elisheva Cohen, and Chandralekha Singh. Learning from mistakes: The effect of students' written self-diagnoses on subsequent problem solving. *The Physics Teacher*, 54(2):87–90, 2016.

[210] Andrew J. Mason and Chandralekha Singh. Impact of guided reflection with peers on the development of effective problem solving strategies and physics learning. *The Physics Teacher*, 54(5):295–299, 2016.

[211] Mark F. Masters and Richard E. Miers. Use of a digital oscilloscope as a spectrum analyzer in the undergraduate laboratory. *American Journal of Physics*, 65(3):254–255, 1997.

[212] Eric Mazur. *Peer Instruction: A User's Manual.* Prentice Hall, Upper Saddle River NJ, 1997.

[213] William F McComas and Michael P Clough. Nature of science in science instruction: Meaning, advocacy, rationales, and recommendations. In *Nature of Science in Science Instruction*, pages 3–22. Springer, 2020.

[214] Lillian C McDermott, Peter S Shaffer, and Mark L Rosenquist. *Physics by Inquiry.* John Wiley & Sons, 1995.

[215] Elissa Milto, Chris Rogers, and Merredith Portsmore. Gender differences in confidence levels, group interactions, and feelings about competition in an introductory robotics course. In *32nd Annual Frontiers in Education*, volume 2. IEEE, 2002.

[216] Özden Karagöz Mirçik and Ahmet Zeki Saka. Virtual laboratory applications in physics teaching. *Canadian Journal of Physics*, 96(7):745–750, 2018.

[217] Paula ML Moya and Hazel Rose Markus. Doing race: An introduction. In Paula ML Moya and Hazel Rose Markus, editors, *Doing Race: 21 Essays for the 21st Century*, chapter 1, pages 1–101. W. W. Norton, 2010.

[218] Forrest S. Mozer and Sondra M. Napell. Instant replay and the graduate teaching assistant. *American Journal of Physics*, 43(3):242–244, 1975.

[219] L. D. Muhlestein and B. DeFacio. Teaching graduate teaching assistants to teach. *American Journal of Physics*, 42(5):384–386, 1974.

[220] National Academies of Sciences, Engineering, and Medicine. *Promising Practices for Addressing the Underrepresentation of Women in Science, Engineering, and Medicine: Opening Doors.* The National Academies Press, Washington, DC, 2020.

[221] National Research Council. *How People Learn: Brain, Mind, Experience, and School: Expanded Edition.* The National Academies Press, Washington, DC, 2000.

[222] National Research Council. *America's Lab Report: Investigations in High School Science.* National Academies Press, 2006.

[223] Knut Neumann and Manuela Welzel. A new labwork course for physics students: Devices, methods and research projects. *Eur. J. Phys.*, 28(3):S61–S69, Apr 2007.

[224] Pasi Nieminen, Antti Savainainen, Niina Nurkka, and Jouni Viiri. An intervention for using multiple representation of mechanics in upper secondary school courses. In Catherine Bruguière, Andrée Tiberghien, and Pierre Clément, editors, *Proceedings of the ESERA 2011 Conference*, Lyon, 2011.

[225] Jayson M. Nissen and Jonathan T. Shemwell. Gender, experience, and self-efficacy in introductory physics. *Phys. Rev. Phys. Educ. Res.*, 12:020105, Aug 2016.

[226] George Allen Brittingham Nock. *The effects on community college student physics achievement and attitudes about learning physics due to inquiry-based laboratory activities versus cookbook laboratory activities*. PhD thesis, University of Mississippi, 2009.

[227] Daniel J O'Brien. A guide for incorporating e-teaching of physics in a post-COVID world. *American Journal of Physics*, 89(4):403–412, 2021.

[228] Valerie K Otero and Danielle Boyd Harlow. Getting started in qualitative physics education research. *Reviews in PER Vol*, 2:1–67, 2009.

[229] Maria Parappilly, Lisa Schmidt, and Samantha De Ritter. Ready to learn physics: A team-based learning model for first year university. *Eur. J. Phys.*, 36(5):055052, Aug 2015.

[230] J Eccles Parsons, TF Adler, R Futterman, SB Goff, CM Kaczala, JL Meece, and C Midgley. Expectancies, values, and academic behaviors. In JT Spence, editor, *Achievement and Achievement Motives*, pages 75–146. Freeman, 1983.

[231] C. Paul and A. Reid. SJSU Real-time instructor observing tool. `http://sjsuriot.appspot.com/`.

[232] Cassandra Paul and Emily West. Using the Real-time Instructor Observing Tool (RIOT) for reflection on teaching practice. *The Physics Teacher*, 56(3):139–143, 2018.

[233] Jean Piaget. *The Origins of Intelligence in Children*. International Universities Press, Inc., New York, 1952.

[234] Anne Marie Porter and Rachel Ivie. Women in physics and astronomy, 2019. Technical report, Statistical Research Center of the American Institute of Physics, 3 2019.

[235] Colin Power. A critical review of science classroom interaction studies. *Studies in Science Education*, 4(1):1–30, 1977.

[236] Katherine N. Quinn, Michelle M. Kelley, Kathryn L. McGill, Emily M. Smith, Zachary Whipps, and N. G. Holmes. Group roles in unstructured labs show inequitable gender divide. *Phys. Rev. Phys. Educ. Res.*, 16:010129, May 2020.

[237] Katherine N. Quinn, Kathryn L. McGill, Michelle M. Kelley, Emily M. Smith, and Natasha Holmes. Who does what now? How physics lab instruction impacts student behaviors. In *Physics Education Research Conference 2018*, Washington, DC, August 1-2 2018.

[238] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2013.

[239] Maureen T Reddy. Invisibility/hypervisibility: The paradox of normative whiteness. *Transformations: The Journal of Inclusive Scholarship and Pedagogy*, 9(2):55–64, 1998.

[240] Edward F. Redish, Jeffery M. Saul, and Richard N. Steinberg. Student expectations in introductory physics. *American Journal of Physics*, 66(3):212–224, 1998.

[241] EF Redish, C Bauer, KL Carleton, TJ Cooke, M Cooper, Catherine Hirshfeld Crouch, BW Dreyfus, BD Geller, J Giannini, J Svoboda Gouvea, et al. NEXUS/Physics: An interdisciplinary repurposing of physics for biologists. *American Journal of Physics*, 82(5):368–377, 2014.

[242] Frederick Reif. *Applying Cognitive Science to Education: Thinking and Learning in Scientific and Other Complex Domains*. MIT Press, 2008.

[243] Muhammad Riaz, Thomas J Marcinkowski, and Ali Faisal. The effects of a DLSCL approach on students conceptual understanding in an undergraduate introductory physics lab. *EURASIA Journal of Mathematics, Science and Technology Education*, 16(2), 2020.

[244] Idaykis Rodriguez, Renee Michelle Goertzen, Eric Brewe, and Laird H. Kramer. Developing a physics expert identity in a biophysics research group. *Phys. Rev. ST Phys. Educ. Res.*, 11:010116, Jun 2015.

[245] Idaykis Rodriguez, Geoff Potvin, and Laird H. Kramer. How gender and reformed introductory physics impacts student success in advanced physics courses and continuation in the physics major. *Phys. Rev. Phys. Educ. Res.*, 12:020118, Aug 2016.

[246] Katemari Rosa and Felicia Moore Mensah. Educational pathways of black women physicists: Stories of experiencing and overcoming obstacles in life. *Phys. Rev. Phys. Educ. Res.*, 12:020113, Aug 2016.

[247] Drew J. Rosen and Angela M. Kelly. Epistemology, socialization, help seeking, and gender-based views in in-person and online, hands-on undergraduate physics laboratories. *Phys. Rev. Phys. Educ. Res.*, 16:020116, Aug 2020.

[248] Laura Ríos, Benjamin Pollard, Dimitri Dounas-Frazer, and H. J. Lewandowski. Pathways to proposing causes for unexpected experimental results. In *Physics Education Research Conference 2018*, PER Conference, Washington, DC, August 1-2 2018.

[249] R. J. Rydell S. L. Beilock and A. R. McConnell. Stereotype threat and working memory: Mechanisms, alleviation, and spillover. *Journal of Experimental Psychology: General*, 136, 2007.

[250] Hannah Sabo, Jennifer Radoff, Andrew Elby, Ayush Gupta, and Chandra Turpen. Role-playing as a tool for helping LAs sense-make about inequitable team dynamics. In *Physics Education Research Conference 2018*, Washington, DC, August 1-2 2018.

[251] Johnny Saldaña. *The Coding Manual for Qualitative Researchers*. Sage, 3 edition, 2016.

[252] Cody Sandifer and Eric Brewe, editors. *Recruiting and Educating Future Physics Teachers: Case Studies and Effective Practices*. Physics Teacher Education Coalition, 2015.

[253] Vashti Sawtelle, Eric Brewe, and Laird H. Kramer. Exploring the relationship between self-efficacy and retention in introductory physics. *Journal of Research in Science Teaching*, 49(9):1096–1121, 2012.

[254] Linda J. Sax, Kathleen J. Lehman, Ramón S. Barthelemy, and Gloria Lim. Women in physics: A comparison to science, technology, engineering, and math education over four decades. *Phys. Rev. Phys. Educ. Res.*, 12:020108, Aug 2016.

[255] Ryan Sayer, Emily Marshman, and Chandralekha Singh. Case study evaluating just-in-time teaching and peer instruction using clickers in a quantum mechanics course. *Phys. Rev. Phys. Educ. Res.*, 12:020133, Oct 2016.

[256] Londa Schiebinger. Has feminism changed science? *Signs*, 25(4):1171–1175, 2000.

[257] Daniel L Schwartz, John D Bransford, and David Sears. Efficiency and innovation in transfer. In J. Mestre, editor, *Transfer of Learning: Research and Perspectives*, pages 1–51. Information Age Publishing, 2005.

[258] Richard M Schwartzstein, Gary C Rosenfeld, Robert Hilborn, Saundra Herndon Oyewole, and Karen Mitchell. Redesigning the MCAT exam: Balancing multiple perspectives. *Acad. Med.*, 88(5):560–567, 2013.

[259] Irving Seidman. *Interviewing as Qualitative Research.* Teachers College Press, 5 edition, 2019.

[260] Elaine Seymour and Nancy M Hewitt. *Talking About Leaving.* Westview Press, Boulder, CO, 1997.

[261] Devyn Shafer, Maggie S. Mahmood, and Tim Stelzer. Impact of broad categorization on statistical results: How underrepresented minority designation can mask the struggles of both asian american and african american students. *Phys. Rev. Phys. Educ. Res.*, 17:010113, Mar 2021.

[262] Manjula D Sharma, Alberto Mendez, Ian M Sefton, and Joe Khachan. Student evaluation of research projects in a first-year physics laboratory. *Eur. J. Phys.*, 35(2):025004, jan 2014.

[263] David J Sheskin. *Handbook of parametric and nonparametric statistical procedures.* CRC Press, 2003.

[264] W-Z Shi, Liping Ma, and Jingying Wang. Effects of inquiry-based teaching on Chinese university students' epistemologies about experimental physics and learning performance. *Journal of Baltic Science Education*, 19(2):289, 2020.

[265] Christopher W. Shubert and Dawn C. Meredith. Stimulated recall interviews for describing pragmatic epistemology. *Phys. Rev. ST Phys. Educ. Res.*, 11:020138, Dec 2015.

[266] C. Singh. Problem solving and learning. In *AIP Conference Proceedings*, volume 1140, pages 183–197, 2009.

[267] Chandralekha Singh. Impact of peer interaction on conceptual test performance. *Am. J. Phys.*, 73(5):446–451, 2005.

[268] Chandralekha Singh. Assessing student expertise in introductory physics with isomorphic problems. I. Performance on nonintuitive problem pair from introductory physics. *Phys. Rev. ST Phys. Educ. Res.*, 4:010104, Mar 2008.

[269] Chandralekha Singh. Assessing student expertise in introductory physics with isomorphic problems. II. Effect of some potential factors on problem solving and transfer. *Phys. Rev. ST Phys. Educ. Res.*, 4:010105, Mar 2008.

[270] Chandralekha Singh. Categorization of problems to assess and improve proficiency as teachers and learners. *American Journal of Physics*, 77(1):73–80, 2009.

[271] Chandralekha Singh. Why flipped classes often flop. *Inside Higher Education*, Jan 2021.

[272] Chandralekha Singh and Guangtian Zhu. Improving students' understanding of quantum mechanics by using peer instruction tools. In *AIP Conference Proceedings*, volume 1413, pages 77–80. American Institute of Physics, 2012.

[273] Ellen A Skinner and Michael J Belmont. Motivation in the classroom: Reciprocal effects of teacher behavior and student engagement across the school year. *Journal of Educational Psychology*, 85(4):571, 1993.

[274] Emily M. Smith, Martin M. Stein, Cole Walsh, and N. G. Holmes. Direct measurement of the impact of teaching experimentation in physics labs. *Phys. Rev. X*, 10:011029, Feb 2020.

[275] Mandana Sobhanzadeh, Calvin S Kalman, and R I Thompson. Labatorials in introductory physics courses. *European Journal of Physics*, 38(6):065702, oct 2017.

[276] David R. Sokoloff. *RealTime Physics: Active Learning Laboratories, Module 4: Light and Optics*. Wiley, 3 edition, 2012.

[277] David R. Sokoloff and Priscilla W. Laws. *RealTime Physics: Active Learning Laboratories, Module 3: Electricity and Magnetism*. Wiley, 3 edition, 2012.

[278] David R Sokoloff, Priscilla W Laws, and Ronald K Thornton. RealTime Physics: Active learning labs transforming the introductory laboratory. *Eur. J. Phys.*, 28(3), 2007.

[279] David R. Sokoloff, Ronald K. Thornton, and Priscilla W. Laws. *RealTime Physics: Active Learning Laboratories, Module 1: Mechanics*. Wiley, 3 edition, 2011.

[280] David R. Sokoloff, Ronald K. Thornton, and Priscilla W. Laws. *RealTime Physics: Active Learning Laboratories, Module 2: Heat and Thermodynamics*. Wiley, 3 edition, 2012.

[281] G. C. Spalding and L. I. McCann. Co-valuing instructional laboratory course offerings. In *Proceedings of the Symposium on Envisioning the Future of Undergraduate STEM Education*, Washington, DC, 2016.

[282] Jacqueline Spears and Dean Zollman. Orientation for the new teaching assistant–A laboratory based program. *American Journal of Physics*, 42(12):1062–1066, 1974.

[283] Jacob T. Stanley and H. J. Lewandowski. Recommendations for the use of notebooks in upper-division physics lab courses. *American Journal of Physics*, 86(1):45–53, 2018.

[284] C. Steele. *Whistling Vivaldi: How stereotypes affect us and what we can do.* WW Norton & Company, 2011.

[285] C. Steele and J. Aronson. Stereotype threat and the intellectual test performance of African Americans. *Attitudes and Social Cognition*, 69, 1995.

[286] Sheldon Stryker. Identity salience and role performance: The relevance of symbolic interaction theory for family research. *J. Marriage Fam.*, 30(4):558–564, 1968.

[287] Lauren L Sullivan, Cissy J Ballen, and Sehoya Cotner. Small group gender ratios impact biology class performance and peer evaluations. *PLOS ONE*, 13(4):e0195129, 2018.

[288] Elli J. Theobald, Melissa Aikens, Sarah Eddy, and Hannah Jordt. Beyond linear regression: A reference for analyzing common data types in discipline based education research. *Phys. Rev. Phys. Educ. Res.*, 15:020110, Jul 2019.

[289] Ronald K. Thornton and David R. Sokoloff. Learning motion concepts using real-time microcomputer-based laboratory tools. *American Journal of Physics*, 58(9):858–867, 1990.

[290] Ronald K. Thornton and David R. Sokoloff. Assessing student learning of Newton's laws: The Force and Motion Conceptual Evaluation and the evaluation of active learning laboratory and lecture curricula. *American Journal of Physics*, 66(4):338–352, 1998.

[291] Adrienne L Traxler, Ximena C Cid, Jennifer Blue, and Ramón Barthelemy. Enriching gender in physics education research: A binary past and a complex future. *Phys. Rev. Phys. Educ. Res.*, 12(2):020114, 2016.

[292] Yaacov Trope and Nira Liberman. Construal-level theory of psychological distance. *Psychol. Rev.*, 117(2):440, 2010.

[293] Chandra Turpen, Ayush Gupta, Jennifer Radoff, Andrew Elby, Hannah Sabo, and Gina Quan. Successes and challenges in supporting undergraduate peer educators to notice and respond to equity considerations within design teams. In *2018 ASEE Annual Conference and Exposition*, Salt Lake City, UT, 07 2018.

[294] David J. Van Domelen and Alan Van Heuvelen. The effects of a concept-construction lab course on FCI performance. *American Journal of Physics*, 70(7):779–780, 2002.

[295] Jonathan B. Velasco, Adam Knedeisen, Dihua Xue, Trisha L. Vickrey, Marytza Abebe, and Marilyne Stains. Characterizing instructional practices in the laboratory: The laboratory observation protocol for undergraduate STEM. *Journal of Chemical Education*, 93(7):1191–1203, 2016.

[296] Trevor S. Volkwyn, Saalih Allie, Andy Buffler, and Fred Lubben. Impact of a conventional introductory laboratory course on the understanding of measurement. *Phys. Rev. ST Phys. Educ. Res.*, 4:010108, May 2008.

[297] Cole Walsh, Katherine N. Quinn, C. Wieman, and N. G. Holmes. Quantifying critical thinking: Development and validation of the physics lab inventory of critical thinking. *Phys. Rev. Phys. Educ. Res.*, 15:010135, May 2019.

[298] Cole Walsh, Martin M. Stein, Ryan Tapping, Emily M. Smith, and N. G. Holmes. Exploring the effects of omitted variable bias in physics education research. *Phys. Rev. Phys. Educ. Res.*, 17:010119, Mar 2021.

[299] Jianye Wei, David F. Treagust, Mauro Mocerino, Anthony D. Lucey, Marjan G. Zadnik, and Euan D. Lindsay. Understanding interactions in face-to-face and remote undergraduate science laboratories: A literature review. *Disciplinary and Interdisciplinary Science Education Research*, 1(14):14, 2019.

[300] Etienne Wenger. *Communities of Practice: Learning, Meaning, and Identity.* Cambridge University Press, 1999.

[301] Etienne Wenger. Communities of practice and social learning systems. *Organization*, 7(2):225–246, 2000.

[302] Candace West and Don H Zimmerman. Doing gender. *Gend. Soc.*, 1(2):125–151, 1987.

[303] Emily A. West, Cassandra A. Paul, David Webb, and Wendell H. Potter. Variation of instructor-student interactions in an introductory interactive physics course. *Phys. Rev. ST Phys. Educ. Res.*, 9:010109, Mar 2013.

[304] Emily A. West, Cassandra A. Paul, David Webb, and Wendell H. Potter. Variation of instructor-student interactions in an introductory interactive physics course. *Phys. Rev. ST Phys. Educ. Res.*, 9:010109, Mar 2013.

[305] Kyle M. Whitcomb, Z. Yasemin Kalender, Timothy J. Nokes-Malach, Christian Schunn, and Chandralekha Singh. Inconsistent gender differences in self-efficacy and performance for engineering majors in physics and other disciplines: A cause for alarm? In *Physics Education Research Conference 2019*, PER Conference, Provo, UT, July 24-25 2019.

[306] Carl Wieman and NG Holmes. Measuring the impact of an instructional laboratory on the learning of introductory physics. *Am. J. Phys.*, 83(11):972–978, 2015.

[307] Carl E Wieman, Georg W Rieger, and Cynthia E Heiner. Physics exams that promote collaborative learning. *The Physics Teacher*, 52(1):51–53, 2014.

[308] B. Wilcox and H. Lewandowski. Students' epistemologies about experimental physics: Validating the Colorado Learning Attitudes about Science Survey for experimental physics. *Phys. Rev. Phys. Educ. Res.*, 12(1):010123, June 2016.

[309] Bethany Wilcox and H. J. Lewandowski. Impact of instructional approach on students' epistemologies about experimental physics. In *Physics Education Research Conference 2016*, PER Conference, pages 388–391, Sacramento, CA, July 20-21 2016.

[310] Bethany Wilcox and Michael Vignal. Recommendations for emergency remote teaching based on the student experience, 2020.

[311] Bethany R. Wilcox and H. J. Lewandowski. Open-ended versus guided laboratory activities: Impact on students' beliefs about experimental physics. *Phys. Rev. Phys. Educ. Res.*, 12:020132, Oct 2016.

[312] Bethany R. Wilcox and H. J. Lewandowski. Developing skills versus reinforcing concepts in physics labs: Insight from a survey of students' beliefs about experimental physics. *Phys. Rev. Phys. Educ. Res.*, 13:010108, Feb 2017.

[313] Bethany R. Wilcox and H. J. Lewandowski. A summary of research-based assessment of students' beliefs about the nature of experimental physics. *Am. J. Phys.*, 86(3):212–219, 2018.

[314] Bethany R Wilcox and HJ Lewandowski. Research-based assessment of students' beliefs about experimental physics: When is gender a factor? *Phys. Rev. Phys. Educ. Res.*, 12(2):020130, 2016.

[315] Matthew Wilcox, Caleb C. Kasprzyk, and Jacquelyn Chini. Observing teaching assistant differences in tutorials and inquiry-based labs. In *Physics Education Research Conference 2015*, pages 371–374, College Park, MD, July 29-30 2015.

[316] Matthew Wilcox, Yuehai Yang, and Jacquelyn J. Chini. Quicker method for assessing influences on teaching assistant buy-in and practices in reformed courses. *Phys. Rev. Phys. Educ. Res.*, 12:020123, Aug 2016.

[317] Edit Yerushalmi, Elisheva Cohen, Andrew Mason, and Chandralekha Singh. What do students do when asked to diagnose their mistakes? Does it help them? I. An atypical quiz context. *Phys. Rev. ST Phys. Educ. Res.*, 8:020109, Sep 2012.

[318] Edit Yerushalmi, Elisheva Cohen, Andrew Mason, and Chandralekha Singh. What do students do when asked to diagnose their mistakes? Does it help them? II. A more typical quiz context. *Phys. Rev. ST Phys. Educ. Res.*, 8:020110, Sep 2012.

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

[320] Barry J Zimmerman. Self-regulated learning and academic achievement: An overview. *Educational psychologist*, 25(1):3–17, 1990.

[321] Benjamin Zwickl, Takako Hirokawa, Noah Finkelstein, and H. J. Lewandowski. Development and results from a survey on students' views of experiments in lab classes and research. In *Physics Education Research Conference 2013*, pages 381–384, Portland, OR, July 17-18 2013.

[322] Benjamin M. Zwickl, Noah Finkelstein, and H. J. Lewandowski. The process of transforming an advanced lab course: Goals, curriculum, and assessments. *American Journal of Physics*, 81(1):63–70, 2013.

[323] Benjamin M. Zwickl, Takako Hirokawa, Noah Finkelstein, and H. J. Lewandowski. Epistemology and expectations survey about experimental physics: Development and initial results. *Phys. Rev. ST Phys. Educ. Res.*, 10:010120, Jun 2014.