Optimizing Mixed Reality Simulation To Support STEM Graduate Teaching Assistants In Developing Student-Centered Pedagogical Skills

Constance Doty
University of Central Florida

Part of the Education Commons, and the Physics Commons
Find similar works at: https://stars.library.ucf.edu/etd2020
University of Central Florida Libraries http://library.ucf.edu

This Doctoral Dissertation (Open Access) is brought to you for free and open access by STARS. It has been accepted for inclusion in Electronic Theses and Dissertations, 2020- by an authorized administrator of STARS. For more information, please contact STARS@ucf.edu.

STARS Citation
OPTIMIZING MIXED-REALITY SIMULATION TO SUPPORT STEM GRADUATE TEACHING ASSISTANTS IN DEVELOPING STUDENT-CENTERED PEDAGOGICAL SKILLS

by

CONSTANCE M. DOTY
B.S. University of Central Florida, Fall 2015
M.S. University of Central Florida, Summer 2018

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Physics in the College of Sciences at the University of Central Florida Orlando, Florida

Fall Term 2021

Major Professor: Jacquelyn J. Chini
ABSTRACT

Physics graduate teaching assistants (GTA) often instruct student-centered lab and recitation sections at large universities, creating the opportunity to positively impact students. However, as STEM GTA professional development varies by institution and discipline, GTAs rarely receive feedback about their teaching. In K-12 teacher preparation, “microteaching”, where a teacher teaches a lesson to peers acting as students, is used to provide feedback. Research has demonstrated benefits to microteaching in K-12 teacher and STEM GTA preparation, but practicing teachers find microteaching can lack authenticity. Recently, researchers have explored integrating technology, like simulation, in pedagogy training to provide a more authentic training experience. However, research is limited on the impact of using simulators for STEM GTA training.

Following a mixed-methods approach, this dissertation presents three studies that focus on the impact of rehearsal in the mixed-reality simulator, TeachLivE™, on physics GTAs’ teaching practices. Study One used observations to investigate GTAs’ use of a questioning strategy both during simulator sessions and while teaching in their actual classroom/lab. We found facilitator feedback was important to GTAs’ rehearsal of questioning during simulator sessions, and we observed an immediate impact on their use of questioning in the classroom. In Study Two, we analyzed student interviews to explore students’ perspectives of GTAs’ use of strategies to resolve group dynamics issues. We found GTAs helped to resolve content-related group challenges but not group dynamics challenges. In Study Three, we analyzed student interviews to
investigate student feelings about GTAs’ use of specific pedagogical skills. We found most students felt nervous if their GTA used cold call, but error framing decreased anxiousness or elicited feelings of support, and most students did not feel interrogated by their GTA’s use of questioning. Overall, these studies demonstrate the utility of simulation to prepare STEM GTAs to use student-centered teaching skills.
ACKNOWLEDGEMENTS

First, I would like to acknowledge my best friend and partner who has always been there for me. Mateo, thank you for all of your support and love all of these years. Thank you for encouraging me to pursue my dreams even though there were some difficult times. Thank you for the sacrifices you have made to help me.

I also would like to acknowledge my advisor, Dr. Jackie Chini (J.C.C.). Thank you for all of your support and patience. Your guidance has helped me to achieve my goals and your encouragement helped me stay motivated. You are the best advisor I could have ever asked for.

Thank you to everyone who collaborated with us on our project. Special thanks to Dr. Erin Saitta (E.K.H.S.) and Dr. Tong Wan (T.W.) who provided numerous support and encouragement. I also would like to thank Ashley Geraets (A.G.G.) and Christopher Nix (C.N.N.), and my fellow PER and CER graduate students.

To my family and friends, thank you for being understanding and encouraging. To my mom, thank you for all that you do for me. I appreciate all of you.

I would like to acknowledge all of my participants. Without you I would not have been able to get this far and write my dissertation. Thank you.

Finally, thank you to all of the wonderful staff members in the physics department, especially Esperanza.

This work was funded by NSF DUE Grant No. 1725554.
# TABLE OF CONTENTS

LIST OF FIGURES ........................................................................................................ xiv

LIST OF TABLES ............................................................................................................ xvi

LIST OF ABBREVIATIONS .............................................................................................. xvii

CHAPTER ONE: INTRODUCTION ............................................................................... 1

CHAPTER TWO: FRAMEWORKS .................................................................................. 5

  GTA Professional Development Evaluation and Design .............................................. 5
  GTA Training Outcomes Impact Students ................................................................. 6
  Factors That Impact GTA Teaching Outcomes ......................................................... 7
  Factors That Impact GTA Training Design .............................................................. 9
  Situated Cognition ...................................................................................................... 9
  Fear of Negative Evaluation ...................................................................................... 11
  Facework .................................................................................................................. 12
  Chapter Summary ..................................................................................................... 14

CHAPTER THREE: LITERATURE REVIEW ............................................................... 16

  GTAs Are Crucial for Undergraduate STEM Learning .............................................. 16
  Active Learning Physics Recitation and Lab Environments ..................................... 17
  Benefits of STEM GTA Professional Development .............................................. 21
  Limitations of Current STEM GTA Training ......................................................... 22
Microteaching Supports Teaching Rehearsal and Feedback, but Might Lack Authenticity ................................................................................................................ 25

Training Simulators Might Provide Ideal Environment for Skill Rehearsal.................. 28

Strategies To Support Instruction in Student-Centered Environments ....................... 29

Student Feelings of Target Skills Need To Be Investigated ....................................... 31

Chapter Summary ...................................................................................................... 33

CHAPTER FOUR: METHODOLOGY ............................................................................ 34

Action Research ......................................................................................................... 34

Case Study ................................................................................................................ 36

Positionality ................................................................................................................ 38

Teaching Intervention: GTA Simulator Training ......................................................... 40

Context ....................................................................................................................... 43

Description of Target Courses ................................................................................ 43

Pre-Existing GTA Training ...................................................................................... 44

Data Collection and Analysis ..................................................................................... 45

Classroom Observation Protocol ............................................................................... 46

Using the Observation Protocol ............................................................................... 46

Pedagogical Skills Added To Protocol .................................................................... 47

Analyzing Classroom Observation Data ..................................................................... 47

Analyzing Student Interview Data .............................................................................. 48
CHAPTER FIVE: RESEARCH QUESTIONS ................................................................. 50

CHAPTER SIX: STUDY ONE – IMPACT OF HIGH-INTENSITY TRAINING WITH A MIXED-REALITY SIMULATOR ON GTAS’ USE OF QUESTIONING ...................... 54

Introduction ................................................................................................................ 54

Frameworks Utilized .................................................................................................. 55

Research Questions .................................................................................................. 55

Methods ..................................................................................................................... 56

Developing GTA Training through Action Research ............................................... 56

Positionality ............................................................................................................ 56

Target Courses – Recitation and Lab Sections ...................................................... 57

Participants ............................................................................................................. 57

Pre-Simulator Training Semester ........................................................................ 57

High-Intensity Simulator Training Semester ........................................................ 58

Pre-existing GTA Training ...................................................................................... 59

Data Collection and Analysis ..................................................................................... 61

Classroom Observations ........................................................................................ 61

Preparing Classroom Observation Data for Comparison ........................................ 64

Sessions with the Mixed-Reality Classroom Simulator ........................................... 64

Findings ..................................................................................................................... 65
Future Work ............................................................................................................................... 85

CHAPTER SEVEN: STUDY TWO – STUDENT PERCEPTIONS OF GTA STRATEGIES TO MANAGE GROUP CHALLENGES ............................................................................................................................... 86

Introduction and Background .................................................................................................. 86

Research Questions ................................................................................................................ 87

Methodology ............................................................................................................................. 88

Using Case Studies to Explore Student Perspective of Group Management ........ 88

Frameworks Utilized ................................................................................................................. 89

GTA Training with a Mixed-Reality Classroom Simulator ......................................................... 89

Participants ................................................................................................................................. 90

Developing Group Challenge Scenarios .................................................................................. 91

Undergraduate Student Interviews ........................................................................................... 92

Data Analysis .............................................................................................................................. 92

Investigating Inter-Rater Reliability ........................................................................................ 93

Findings and Discussion ............................................................................................................ 93

Students Experienced Group Challenges Described by the Scenario Prompts.................................. 94

GTAs Might Not Recognize Group Dynamic Challenges ......................................................... 98

Students Experienced Alternative Group Challenges .......................................................... 101

Student Frustration with Tutorials Might Influence Group Members to Not Contribute .................. 102
Students Perceived GTAs to Help with Group Challenges Related to Physics

Student Perspective of GTA's Role: Mentor vs. Instructor

Mentors Create Comfortable Learning Environments

GTAs Perceived as Mentors Help Resolve Group Challenges

Implications for GTA Preparation

GTAs might recognize group challenges but choose not to engage

Conclusion

Limitations

Future Work

CHAPTER EIGHT: STUDY THREE – IMPACT OF TARGET SKILLS ON STUDENT FEELINGS

Introduction and Background

Research Questions

Methodology

Using Case Studies to Explore Student Perspective of Pedagogical Skills

Frameworks Utilized

Participants

Undergraduate Student Interviews

Data Analysis
LIST OF FIGURES

Figure 1: Adapted framework representation from Reeves et al., (2016) ...................... 6
Figure 2: A Visual Representation of Connection Between Frameworks and Studies .. 15
Figure 3: Picture of a researcher (T. W.) interacting with the simulator. Left Image:
Virtual classroom projected on a screen in front of simulator user; Right Image:
Researcher (T. W.) interacting with simulator in a physical room ............................ 41
Figure 4: Schematic of the Cyclical Procedure for a Simulator Session .................. 42
Figure 5: Timeline of all semesters. A – Pre-simulator Semester; B – Simulator Training
Semester 1; C – Simulator Training Semester 2 .......................................................... 45
Figure 6: Timeline for Both Cohorts. A) (Top) Timeline of Fall 2018 Pre-Simulator
Intervention. B) Timeline of Fall 2019 High-intensity Simulator Intervention .......... 63
Figure 7: GTAs’ Use of Stretch-It during Questioning Skills Session: (A) (Left): Round 1;
Analogous Situation”, AW – “Ask for Another Way to Answer” .............................. 66
Figure 8: Stretch-It Count across Simulator Sessions. (A) Session One; (B) Session
Two; (C) Session Three; (D) Session Four. The four sub-categories of Stretch-It are as
Situation”, AW – “Ask for Another Way to Answer” ................................................. 68
Figure 9: High-intensity Simulator Training (Fall 2019) Cohort’s Average Use of Explain
Logic. Each point on the chart represents the average count of two-minute intervals
Explain Logic was coded across GTAs who were observed teaching that lesson. Both
lines represent the average count of two-minute intervals Explain Logic was coded
across observations for pre-simulator (Fall 2018) and high-intensity training (Fall 2019). Shading represents the standard error of the mean (as error bars) for each semester average. ........................................................................................................................ 71

Figure 10: High-intensity Simulator Training (Fall 2019) Cohort’s Average Use of Explain Logic without Outlier Observation. Each point on the chart represents the average count of two-minute intervals Explain Logic was coded across GTAs who were observed teaching that lesson. Both lines represent the average count of two-minute intervals Explain Logic was coded across observations for pre-simulator training (Fall 2018) and high-intensity simulator training (Fall 2019) semesters. Fall 2018’s average does not include the outlier observation. Shading represents the standard error of the mean (as error bars) for each semester average. ......................................................... 73
LIST OF TABLES

Table 1: Pre-Simulator Training Semester Participants ................................................. 57
Table 2: High-intensity Simulator Training Semester Participants ................................. 58
Table 3: Questioning Skills Presented in Simulator Intervention ................................... 60
Table 4: GTAs’ Use of Stretch-It: Explain Logic in the Real Classroom ......................... 75
Table 5: Undergraduate Physics Student Participants .................................................. 90
Table 6: Students Report Experiences Similar to the Prompted Group Scenarios ....... 95
Table 7: Students Report Group Challenges Not Described by the Prompted Group Scenarios ................................................................................................................. 104
Table 8: Relating Experience with Group Challenges And Perception of GTA ............ 107
Table 9: Undergraduate Student Participants ............................................................. 118
Table 10: GTA Responses to Incorrect Student Idea Descriptions ............................. 126
Table 11: GTAs’ Use of Stretch-It: Follow-Up in the Real Classroom .......................... 166
LIST OF ABBREVIATIONS

1o1_TPQ – One-on-One TA Posing Question

BER - Biology Education Research

CER – Chemistry Education Research

DBER – Discipline-Based Education Research

FNE – Fear of Negative Evaluation

GTA – Graduate Teaching Assistant

IRR – Inter-Rater Reliability

ISLE – Investigative Science Learning Environment

LA – Learning Assistant

LOPUS – Laboratory Observation Protocol for Undergraduate STEM

MILE - Mixed-reality Integrated Learning Environment

PER – Physics Education Research

PQ – (TA) Posing Question

RIOT – Real-time Instructor Observation Tool

STEM – Science Technology Engineering and Mathematics

Stretch-It: Explain Logic – Explain Logic

Stretch-It: Follow-Up – Follow-Up
TA – Teaching Assistant

TLE TeachLiVE™ - TeachLivE
CHAPTER ONE: INTRODUCTION

In response to a national call to improve STEM education by implementing instruction that supports student engagement during the learning process [1] and the reported benefits of active learning instruction for student learning [2], many physics departments across the United States have adopted active learning student-centered instruction in recitation and lab sections [i.e., 3-9]. Graduate Teaching Assistants (GTA) are often appointed the responsibility of leading such courses, and therefore, are crucial to the success of postsecondary STEM education. For example, GTA-initiated interactions in physics labs and GTAs’ use of questioning instead of confirmation during tutorials have been linked to enhanced student learning [10,11]. However, GTAs in such spaces have been observed to use a variety of teaching styles, some of which are not complementary to student-centered instruction. For instance, West et al. reported physics lab teaching assistants to vary in their teaching behaviors, including higher percentages of non-interactive behaviors (e.g., waiting or observing) than expected by the course designer [12]. Similarly at UCF, Wilcox, Yang and Chini found physics GTAs who taught mini-studio (combined recitation and lab) sections to have “buy-in” associated with the student-centered curricula, but their teaching practices did not align with their teaching beliefs as they taught how they thought their students wanted them to teach [3]. It is possible GTAs need scaffolding to develop pedagogical skills that support such instruction since there are limited opportunities for GTAs to receive feedback about their teaching or practice using teaching skills before entering the classroom.
Microteaching, also referred to as role-playing, is a successful training model used in K-12 teacher preparation. In microteaching teacher trainees take turns teaching a lesson while their peers “play” students [13]. After the lesson, they receive feedback from their peers and facilitator, similar to a deliberate practice model [12]. Beginner pre-service teachers have reported microteaching led to feeling prepared to teach in their actual classrooms [15] and have expressed feedback during microteaching lessons helps to improve their teaching skills [16]. For these reasons, the microteaching lesson model is currently being explored in STEM GTA preparation [17,18]. Doucette et al. found physics lab GTAs engaged in a higher percentage of open dialogue with students than GTAs who did not participate in microteaching sessions [18]. In biology, microteaching is also being explored to scaffold teaching assistants’ use of pedagogical skills. Becker et al. found a positive impact of such training on teaching assistants’ use of some of the skills in the classroom, but the impact was not stable [19]. Therefore, further exploration of training that uses microteaching is necessary to inform future GTA professional development.

A limitation to microteaching is using GTAs to portray students as they might not exhibit authentic student behaviors (supported by situated cognition [20]). An alternative to microteaching is to incorporate sessions with the mixed-reality classroom simulator, TLE TeachLivETM, into current STEM GTA professional development. The simulator was originally developed for pre-service teachers to practice teaching to a virtual class in a safe and authentic environment before teaching in a real classroom with real students [21,22]. Since its development, the simulator has been used for various applications, like teacher practice with an avatar-student with autism [23] and providing
a space for law enforcement to practice de-escalation strategies [24]. In addition, the simulator has been used as a tool for training physics Learning Assistants (LA). Chini, Straub, and Thomas (2016) found six out of nine physics LAs improved at least one pedagogical skill mentioned in their personal teaching reflections while rehearsing in the simulator [25]. However, research about using a simulator as a tool for physics GTA training is limited. Another study about a different mixed-reality simulator, mixed-reality integrated learning environment (MILE), investigated the GTAs’ performance in the simulator and teaching self-efficacy [26], but did not measure the impact of the simulator training on GTAs’ performance in their actual classrooms. Therefore, this dissertation will contribute to the existing literature about GTA professional development as well as potential applications of the use of mixed-reality simulators in teacher preparation.

In this dissertation, three studies exploring the impact of the simulator training on GTAs and students will be presented. The first study explores GTAs’ use of higher order questioning (operationalized as “Stretch-It”) during simulator rehearsal sessions and in the classroom. The second study investigates the student perspective to explore how GTAs notice and respond to group dynamics issues in physics mini-studio sections. The third and final study investigates students’ perspective about their GTA’s use of the target skills and impact on their feelings of anxiousness and learning environment.

The results of this dissertation will notify the PER and DBER communities about the potential impact of implementing training sessions with the mixed-reality simulator, TeachLivE™, into training on GTA and student outcomes. Following a positive impact of the training on GTA teaching behaviors and student outcomes, we would recommend
the integration of a deliberate practice model into STEM GTA training along with the use of technology.
CHAPTER TWO: FRAMEWORKS

In this dissertation, we used four frameworks to guide the study design (e.g., teaching intervention, data collected, methods of analysis) and the interpretation of our findings (e.g., implications, discussion). By using the included frameworks, we ensure that we are in alignment with best practices for developing training for STEM GTAs and investigating the impact of such training on GTA and student outcomes. Here, we discuss an encompassing conceptual framework called GTA Professional Development Evaluation and Design. Then, we discuss details about the situated cognition framework which motivated pieces of the teaching intervention design. Next, we present fear of negative evaluation (FNE) which informed us about a one possible reason students may hesitate to share ideas with their peers. Finally, we discuss facework, a framework we used to operationalize the actions GTAs might take to mitigate the negative impact of FNE.

GTA Professional Development Evaluation and Design

The model of GTA training design and assessment presented in this dissertation was informed by a conceptual framework created by Reeves et al. focused on assisting GTA training facilitators and researchers with assessing and designing GTA professional development programs [27]. The framework highlights connections between three constructs of GTA preparation: desired training outcomes (GTA and student), context of the training program (institutional factors and training design), and moderating factors (implementation of training and GTA characteristics). A visual representation is provided in Figure 1.
GTA Training Outcomes Impact Students

Desired outcome variables of the GTA training framework are aligned with the initiatives and goals stated by the Biology Teaching Assistant Project (BioTAP) [28], such that facilitators and researchers focus attention on measuring changes in GTA cognition (e.g., knowledge, skills, and beliefs about teaching and student learning), GTA instructional practices, and student outcomes (e.g., learning, retention, affective behaviors, etc.). The framework suggests that student outcomes are influenced by GTA teaching variables (GTA cognition and instructional practices). For example, if a GTA implements student-centered instruction in the lab or recitation, students might perform better in the course than if the GTA used traditional methods, as suggested by Freeman et al [2]. Furthermore, instructional practices used by GTAs are influenced by GTA cognition factors. If a GTA agrees with a teacher-centered approach to learning, then their teaching style might include behaviors like explaining at the board rather than engaging in open dialogue with students. Training programs can influence pieces of
GTA cognition, like knowledge of strategies to encourage students to share their ideas. For instance, Lee reported new GTAs shifted their teacher-centered philosophy to be more aligned with student-centered ideas after completing a pedagogy course [29]. Additionally, GTA training can impact GTA teaching practices, like development of pedagogical skills used by GTAs in their classroom. Becker et al., applied this framework when implementing new training for biology GTAs and found GTAs used skills from the training in their classrooms [19]. Since we aim to support GTAs’ use of pedagogical skills in the classroom which impact students, we use this framework to inform us about factors that impact GTAs’ classroom practices.

Factors That Impact GTA Teaching Outcomes

The framework suggests training design and implementation in parallel with GTA characteristics influence GTA outcomes. Since we aim to impact GTA teaching practices used in the classroom, we designed our training program to include content and activities that help with GTAs’ development of pedagogical skills complementary to student-centered instruction. The nature of the training activity is supported by using another framework, situated cognition (described in detail later in this section), and examples of such activities are included in the Literature Review Chapter.

Implementation details of such training are important to consider when reporting the impact of the teaching intervention on GTA outcomes. Two examples include fidelity of implementation by facilitators and GTA participation in the training. If facilitators do not implement the training as intended by designers, then the training might not have the intended impact on GTA outcomes. Research has similarly shown that faculty
instructors may leave out important parts of student-centered activities developed by STEM education researchers to benefit student learning [30-32], impacting their students’ outcomes. Moreover, if GTAs do not engage in the training as intended by designers, like not reading through training material before attending training sessions, then the GTAs might not engage with the training as intended by designers. In detail, GTAs’ responsiveness to the training (e.g., Do they rehearse the skills in the simulator?) and the degree to which trainers align the facilitation of training with the design might influence GTAs’ transfer of training aspects to the classroom. From this reasoning, we investigate GTAs’ use of the skills in the simulator sessions and possible reasons why they might not use the skill in the classroom (within discussion sections of Studies One through Three) during the feedback portion of the training.

Additionally, GTAs have a variety of previous teaching experiences and ideas of teaching that influence their instructional practices. For example, it is possible GTAs might neglect to include a piece of student-centered instruction for which they do not agree. For example, Goertzen, Scherr, and Elby presented a case study of a physics tutorial GTA named “Oscar” [33]. Oscar did not agree with the portion of the tutorial that asks students to relate their common sense to the physics situation because Oscar did not believe that portion was beneficial for student learning, leading Oscar to ignore asking students about that portion of the tutorial [33]. Also, GTAs might differ in their conceptualization of the pedagogical skills from training designers and education literature. An example of this is included in Gereats’ dissertation which reported chemistry GTAs (part of the larger project) conceptualize error framing (e.g., framing a student mistake as natural or beneficial to learning) in different ways than the training
facilitators [34]. Therefore, Reeves et al. advise facilitators and researchers to explore GTA characteristics throughout the process of training design and when reporting findings [27]. In our analysis, we consider GTAs' teaching experience to be a variable when reporting our findings. We also recognize GTA characteristics as a limitation to the study as we cannot control for GTA past experiences in education.

Factors That Impact GTA Training Design

Since GTA training initiatives are embedded in institutional contexts like institution type and student enrollment (size and demographics), it is necessary to consider how institutional factors might influence GTA training design. At research-intensive universities, departmental culture might propagate the norm of favoring research over teaching [35]. However, members of the physics education community have argued that physics departments should support teaching as a proper career [36] since graduate students often become future faculty members with teaching responsibilities [37]. Therefore, we use this piece as a call to action for members of the physics (and chemistry) department to help create a culture supportive of reformed teaching practices for GTAs.

Situated Cognition

Situated cognition describes learning as a context-based process such that theory and practice of knowledge are not separable [20]. Practitioners and members of the community shape the purpose and use of discipline-based knowledge in such situations. Extending this framing to education, teachers' beliefs and practices
associated with teaching and learning are situated in classroom environments. Both beliefs and practices can be influenced by the teaching and discipline-specific communities. For example, postsecondary physics instructors, like physics GTAs, can choose to use evidence-based teaching practices investigated and suggested by researchers from the discipline-based education research (DBER) group in their classrooms. However, physics GTAs might not find the use of evidence-based practices from DBER literature to be intuitive to apply in their classes. For example, Wilcox, Yang and Chini, observed physics GTAs to misalign their teaching practices with their beliefs, despite participating in a pedagogy seminar and having “buy-in” for the student-centered curricula [3].

Microteaching applies the model of situated cognition to teacher preparation. In microteaching sessions, participants alternate between the role of teacher and student while playing out classroom situations [13]. When playing the role of teacher, participants can rehearse putting pedagogy into practice before using the same strategies in their real classroom. When paired with the concept of deliberate practice [14], participants can rehearse teaching, receive feedback from peers and facilitators, and try the same teaching scenario again with new goals. However, participants playing students might not simulate authentic student reactions or feelings. To mimic a more authentic classroom experience, classroom simulators, like the TeachLivE mixed-reality classroom simulator, can be used instead.

TeachLivE™ (referred to as the simulator in later chapters) was developed for pre-service teachers to rehearse and receive feedback about their interactions with
simulated students (avatar-students) before teaching real students in a real classroom [21]. The simulator combines aspects of mixed-reality (participant interacts with a virtual classroom projected on a screen in a physical room) and human-in-the-loop technology to create an authentic experience while maintaining brevity for training purposes. Most simulator users have reported forgetting that they were interacting with avatar-students in a simulator because the interactions were like talking with real people or students [22,25]. Participants’ suspension of disbelief can be leveraged for short periods of time (ten minutes or less) to create an authentic teaching experience for novice teachers to rehearse pedagogical skills without harming real students.

Since we implemented simulator sessions focused on the rehearsal of pedagogical skills into physics GTA training, we apply the model of situated cognition to operationalize the cognitive process of GTAs participating in the simulator training presented in this dissertation.

**Fear of Negative Evaluation**

In active learning environments, there are more opportunities for student-student and student-instructor interactions to occur compared to traditional courses. When students share their ideas, they might create a social situation where they are subjected to evaluation. If students share incorrect or unexpected ideas, they might feel anxious when anticipating or receiving negative judgement about their idea from their peers and/or instructor. This feeling might discourage students from wanting to further participate in activities where they are expected to interact with their peers, like group work. Watson and Friend referred to this as fear of negative evaluation (FNE) [38]. More
specifically, they defined FNE as, “apprehension about others' evaluations, distress over their negative evaluations, avoidance of evaluative situations, and the expectation that others would evaluate oneself negatively” [38, p. 449].

Furthermore, instructors in active learning courses might choose to implement evidence-based strategies to encourage students to participate in class, but they might not realize the impact the strategy has on student affect. For example, the use of cold call (i.e., when students are called on by name to share their ideas in front of their peers without volunteering) by instructors has been linked with an increase in student participation (i.e., more students wanting to volunteer their answers) [39]. However, Cooper, Downing and Brownell found 59.6% of biology undergraduate students they interviewed reported increased levels of anxiety mostly attributed to FNE when their instructor used cold call during lecture; no students in their sample mentioned a decrease in anxiety associated with cold call [40].

Since we encouraged GTAs to rehearse and use evidence-based strategies in their teaching, like cold call, it is important to investigate the impact of their use of pedagogical strategies on student feelings. Therefore, in Study Three, we used the concept of FNE to motivate undergraduate student interviews enrolled in sections where their GTA used the target strategies from the simulator training.

Facework

Since the nature of active learning instruction encourages the sharing of ideas between students and their instructor, an active learning classroom is a socially
dynamic space, and such social interactions can impact the social image of each person involved. Goffman referred to a person’s preferred self-image that they wish to preserve as *face* [41]. Connected to the previous framework, a student’s face can be threatened by fear of negative evaluation. For example, if a student shares an incorrect idea in front of their peers, they might internalize the belief that their peers or instructor think that they are stupid.

Furthermore, an instructor can take social actions to maintain and prevent negative threats to a student's face, also referred to as *facework* [41]. An instructor’s use of protective facework (i.e., working to protect or preserve a student’s face) might help create an environment supportive of student learning. For example, Gaffney and Gaffney found a correlation between facework and student satisfaction in their physics course [42]. Kerssen-Griep, Hess and Trees found a link between facework and student motivation [43].

The use of student names when interacting with students [44] and discussions about preconceptions of physics on the first day of class are examples of facework that instructors can use in the classroom [45]. When using student names, students might feel as though their instructor cares about them as an individual, breaking down the power dynamic in the classroom [44]. Discussions about students' initial ideas about physics might help them feel as though their instructor is supportive of their learning [45]. Additionally, teaching strategies that frame mistakes as natural or beneficial to the learning process, such as error framing [46], might mitigate the negative social threats caused by FNE on a student’s face when they share an incorrect answer. For example,
Downing et al. found community college students enrolled in active learning biology courses expressed their instructor's use of error framing reduced their anxiety when they participated in a whole class discussion [47]. In our previous study, we interviewed eleven students enrolled in either a student-centered chemistry or physics lab about their GTA's use of cold call. We found most students reported their GTA to pair cold call with another teaching strategy (including error framing) that helped reduce the anxiety associated with being called on to share an idea in front of the class [48]. From this preliminary finding, we argue that GTAs can use facework to create an environment where students feel comfortable sharing their ideas, and we continue our investigation in Study Three where we operationalized the teaching strategies used by GTAs to create a comfortable learning environment as protective facework.

Chapter Summary

As described above, we used four frameworks in this dissertation and described the connections between the frameworks and the studies (a visual representation is provided in Figure 2). All three studies were informed by the framework detailed by Reeves et al., including GTA and student data to collect, how to analyze the data, and how we interpret the findings. In Study One, we applied the Reeves et al. framework to assess the impact of training design on GTA outcomes. Additionally, in Study One, we used the situated cognition framework as motivation for the use of a classroom simulator paired with deliberate practice for the GTA training activity (teaching intervention).
Since the Reeves et al. framework informs us about the impact of GTA teaching outcomes on students, we focus our attention on the connection between GTA teaching outcomes and students in their classroom. Additionally, we recognize students might not feel comfortable sharing their ideas with their peers and the pedagogical skills presented in the teaching intervention might specific student feelings. Thus, in Studies Two and Three, we operationalized students’ fear of sharing ideas as the concept of fear of negative evaluation.

Since we are investigating the connection between GTA and students, the Reeves et al. framework suggests a GTA’s action in the classroom impact student outcomes. Therefore, in Study Two and Three, we applied the Facework framework to describe actions GTAs take to help students feel comfortable sharing their ideas.
CHAPTER THREE: LITERATURE REVIEW

In this chapter, we provide motivation and background information crucial for investigating STEM GTA training. Additionally, we describe how this dissertation contributes to the literature focused on physics GTA training by identifying places where further investigation is necessary. First, we introduce GTAs as postsecondary STEM instructors. We also describe the interactive classroom environments in which physics GTAs often teach, including how students feel in these spaces. Then, we discuss the benefits of GTA training and identify limitations of current STEM GTA professional development. Informed by K-12 teacher preparation, we describe microteaching as a potential training activity to support GTAs’ development of pedagogical skills through of practice and feedback. Given the limitations of microteaching, we propose using a mixed-reality classroom simulator to create a more authentic rehearsal experience for GTAs.

GTAs Are Crucial for Undergraduate STEM Learning

Motivated by initiatives to transform the STEM classroom into an interactive learning environment [1] and the reported benefits of active learning instruction on student learning [2,49], postsecondary STEM departments might choose to adopt an active learning model as a replacement for traditional-style instruction in recitation and lab sections. Several physics departments have replaced traditional-style recitations with tutorials [3-7] and some have switched from implementing confirmation labs to inquiry or process-focused curricula [3,8,9]. Given high student enrollment in postsecondary introductory STEM courses, GTAs are often given responsibility for
teaching sections of recitation (tutorial) and lab [35]. In our previous study, we found a
correlation between GTA instructional style in introductory mechanics mini-studio
sections and post-test score on the Force Conceptual Inventory with small effect size
[50]. Thus, STEM GTAs are crucial to student learning experience.

Furthermore, student perceptions of GTAs can help foster a comfortable learning
environment for students. Muzaka found students perceived their biology GTA to be
more approachable and better at facilitating group discussions than faculty members
[51]. Similarly, Kendall and Schussler reported students perceived their GTA to be more
relatable than faculty instructors [52]. However, both studies reported students to
perceive their GTA to have novice teacher traits, such as uncertainty with teaching
practices [52] and having knowledge specific to the lab course only [51]. Since GTAs
have an important responsibility to support student learning, GTA training can
incorporate elements to help GTAs become more confident in implementing evidence-
based teaching practices while thinking about how to support student learning within the
entirety of the course.

Active Learning Physics Recitation and Lab Environments

Driessen, Knight, Smith, and Ballen have described active learning as "an
interactive and engaging process for students that may be implemented through the
employment of strategies that involve metacognition, discussion, group work, formative
assessment, practicing core competencies, live-action visuals, conceptual class design,
worksheets, and/or games." [53]. Thus, in an (student-centered) active learning
environment, students are expected to actively participate in class while engaging in
activities that encourage them to think about, discuss and share their ideas with their instructor (GTA) and peers. Within this instructional model, knowledge can be “constructed” through social situations like conversations about conceptual understanding with group mates (i.e., social constructivism [54]). Thus, an active learning classroom (or lab) is full of opportunities for students to interact with their peers and their GTA, and those interactions might impact student learning experience.

Participants in this dissertation taught or were enrolled in mini-studio sections associated with either portion of a two-semester introductory algebra-based physics course. Mini-studio mimics a studio-mode course on a smaller scale and combines two different active learning activities, tutorial-style recitation (tutorial) and lab, into one three-hour weekly session. Within each portion of mini-studio, there were more opportunities for a student to interact with their instructor (GTA) and peers compared to lecture, creating an ideal space to implement active learning instruction.

During tutorial, students worked together in small groups to complete a worksheet that shifts students’ goal of getting the correct answer to making sense of physics or “sensemaking” (i.e., Tutorials on Physics Sensemaking [7]). Students were expected to share and discuss their own understanding of the physics phenomenon while checking for alignment with everyday thinking within their group. In line with a call to shift the focus on physics labs from reinforcing lecture concepts to developing experimental skills [9, 55], students were expected to work together on experiments designed to emphasize the process of scientific investigation rather than confirming a
theoretical model from lecture or the textbook (i.e., Investigative Science Learning Environment referred to as ISLE, [8]).

Since active learning suggests a connection between socialization and learning, GTA-student interactions might positively impact a student’s learning experience. During an initial investigation of GTA behaviors during tutorial, Scherr et al. found GTAs initiated more interactions with students than expected by curriculum designers [56]. In a different study, Goertzen, Scherr, and Elby observed tutorial GTAs to use a variety of strategies to “look for indicators” of student understanding [57]. When investigating GTA-student interactions in an introductory physics lab, Stang and Roll found a positive correlation between frequency of GTA-initiated interactions with students and improved student engagement, a construct linked with student learning [10]. These findings suggest GTAs play a role in creating an environment supportive of student learning.

However, physics GTAs have been observed to inconsistently implement the same tutorial or lab curriculum. To investigate GTA teaching practices in the classroom, West et al. created the real-time instructor observation tool (RIOT) to measure the context of teaching assistant-student interactions during an interactive course lead mostly by GTAs [12]. They found the instructors of the course held a variety of instructional styles, including a less interactive teaching style not intended for such an environment [12]. Wilcox, Yang, and Chini applied the same protocol to explore GTA-student interactions in introductory physics mini-studio (combined tutorial and lab) sections [3]. While most GTAs had “buy-in” for student-centered instruction, they found GTA teaching practices did not align with their teaching beliefs [3]. This finding is not
unique to physics. Goodwin et al. found biology GTAs value and are knowledgeable about evidence-based teaching practices, but their teaching beliefs were not reflected in their teaching practices [58]. Since GTAs play a role in student learning, it is crucial to provide GTAs with training to support their development of evidence-based practices in their classroom or lab.

Furthermore, Conlin and Scherr argued a safe space to “sensemake” must be established within the student group to achieve the optimal benefits of tutorials on student learning [59]. In their comparative case study, Conlin and Scherr investigated student behaviors during the formation of a safe space during the first session of tutorial [59]. They found two out of three student groups successfully created a space to sensemake using epistemic distancing (e.g., communicative behaviors that soften a student’s stance) [59]; however, one group was not successful in creating such a space. Similarly, in lab spaces, it is crucial for students to create a safe space to productively reason through the inquiry process (i.e., “problematizing”) [60]. Brookes, Yang, and Nainabasti investigated social positioning (e.g., a student’s stance as expert or novice, and soft or firm) of students in ten different groups during ISLE lab sessions [61]. They found group members equally contributing to the group discussion resulted in the formation of a more effective group (e.g., student group discussed and completed activity learning objectives) than groups that were less equitable [61]. Also, in their example of an equitable group, they observed that one group with equitable contributions identified and worked together to solve an error rather than giving up on the task like other groups [61]. Both studies imply social dynamics of the group might influence student engagement during tutorials and inquiry-style labs. Since GTA-student
interactions also influence student engagement in both settings, it is possible GTAs might play a role in a student group’s formation of a safe space. However, literature documenting strategies used by GTAs to help create such a space is limited. Thus, it is important to investigate how GTAs help students create these spaces and resolve group dynamic challenges.

Benefits of STEM GTA Professional Development

Physics GTAs hold three main roles in a department: teacher, student, and researcher. However, GTAs might struggle with their self-worth and identity in their department when balancing each role [62]. Their struggle might partly be due to the de-emphasis of teaching as a useful STEM identity. In chemistry, Schussler et al. found GTAs reported an absence of faculty members valuing graduate students spending time on teaching-related professional development as a barrier to participating in GTA training [63]. Similarly in physics, Luft et al. discussed how faculty mentors and GTAs might not place enough emphasis on teaching because it is a secondary role as a scientist [35]. Sandi-Urena and Gatlin found GTA training to have a positive impact when STEM GTAs are building their teaching identity [64]. Reeves et al. investigated professional development offered at three different institutions in the United States and found GTAs who received any form of training reported an increase in teaching self-efficacy and pedagogical content knowledge and a decrease in anxiety associated with teaching [65]. Therefore, professional development can provide STEM graduate students with opportunities to foster a healthy teaching identity.
Furthermore, developing pedagogical skills during graduate school can help future faculty members become proficient in implementing student-centered instruction [37]. Brownell and Tanner found most biology faculty members might not be prepared to implement student-centered instruction [66]. Stains and Vickrey found STEM faculty members struggle to implement evidence-based practices which can lead to a lesser impact on student learning [31]. Therefore, implementing training with evidence-based practices and student-centered instruction earlier in a faculty member’s career, like graduate school, is beneficial for faculty development and student learning.

Additionally, GTAs who teach inquiry-style labs have reported benefits from their experience as an instructor on their research design skills [67,68]. As previously reported, GTAs have been observed to inconsistently implement such curriculum. Therefore, it is important to encourage GTAs to participate in teacher training to align teaching beliefs and practices for optimal benefits.

Limitations of Current STEM GTA Training

According to a report from the American Institute of Physics (AIP), 72% of first-year physics graduate students received teacher preparation between 2014 and 2016, but the training activities GTAs participated in were not specified [69]. Typically, STEM GTAs are provided a variety of training opportunities at the university and department-levels [35,70]. Here, we discuss examples of STEM GTA training at both levels.

University-wide training is often a mandatory workshop offered to GTAs in their first semester of their teaching assignment and occurs before GTAs begin teaching
The content of the workshop is often generic and focuses on topics such as university policies. Mandatory university training is not continued throughout a GTA’s teaching assignment. There are opportunities for graduate students to become involved in cross-institutional efforts (including their own institution), like the Preparing Tomorrow’s Faculty program. However, like other university-provided training, the topics are aimed at a wider audience and might not cover the applications of teaching in a specific discipline. It is possible STEM GTAs might not find all the topics in such a program relevant to their discipline.

The physics education community has called on physics departments to support and provide resources for teacher preparation. For example, Otero, Pollock, and Finkelstein created the Learning Assistant Program to encourage and prepare undergraduate physics students to take on teaching roles after they graduate. Since its creation, the program has been adopted by physics departments in other institutions and within other STEM disciplines. While a national program does not currently exist for GTA preparation, STEM education literature has documented examples of GTA training activities offered at the departmental level, like front-loaded workshops, pedagogy seminars, and weekly preparation meetings, described in detail below.

Front-loaded workshops provide initial scaffolding for GTAs to get started with teaching, but continued training is necessary for supporting GTAs while they teach. For example, Pentacost et al. developed a three-day course to introduce incoming chemistry GTAs to the newly reformed general chemistry recitation by modeling the
expected learning environment [78]. While participating GTAs found the training to be useful because they had an opportunity to model the expected learning environment, they also wished the training continued throughout the semester to help them implement student-centered instruction (the second iteration of the training included meetings throughout the semester) [78].

Unlike beginning of the semester workshops, weekly preparation meetings and pedagogy seminars provide GTAs with longer term support. During preparation meetings, GTAs might have opportunities to work through the next week’s activity as a group and discuss common student ideas about learning within the context of the activity, as suggested by McDermott and Shaffer for physics tutorial GTAs [79]. Due to barriers, like time constraints, such meetings might focus on content rather than classroom discourse [71]. Pedagogy seminars provide a space for GTAs to develop pedagogical content knowledge (PCK) vital for teaching [80]. In such courses, GTAs read and reflect upon articles about teaching pedagogy, participate in group discussions of active learning strategies, and create an active learning activity [29]. However, it might not be intuitive for GTAs to put pedagogy into practice. For example, Wilcox, Yang and Chini observed physics GTAs to misalign their practices with their agreement with student-centered instruction despite participating in a pedagogy seminar and weekly preparation meetings [3].

One reason might be due to the lack of opportunities for rehearsal of evidence-based practices and feedback about their teaching in current STEM GTA training efforts. Of the biology GTAs surveyed by Goodwin, Cao, Fletcher, Flaiban and
Shortlidge, 72% of GTAs mentioned misalignment between teaching beliefs and practices might be due to the absence of training to scaffold their use of evidence-based practices in the classroom [58]. Thus, STEM GTA professional development should create opportunities to scaffold the use of evidence-based practices with rehearsal and feedback.

**Microteaching Supports Teaching Rehearsal and Feedback, but Might Lack Authenticity**

Speer, Guttman, and Murphy have argued for STEM GTA professional development to incorporate models developed for K–12 teacher preparation because the success of such programs on beginning teacher’s instructional practices are well documented compared to GTA training programs [81]. Microteaching is a training activity used in K–12 teacher preparations similar to deliberate practice (i.e., training activities provide opportunities for intentional practice and are iterated for expert-like proficiency) [14]. In microteaching, teachers alternate between teaching a lesson to their peers and “playing” students; afterwards, facilitators and peers provide feedback about their teaching performance [13]. Education research has documented reported benefits of microteaching sessions for K-12 student teachers like increased teaching self-efficacy [82] and increased knowledge about reform-based teaching strategies [83]. Additionally, incorporating microteaching sessions into STEM GTA training might provide GTAs with opportunities to receive feedback about their teaching, which K–12 student teachers have expressed helped them feel prepared for teaching their actual students [15] and pinpoint their strengths and weaknesses to further pedagogical skill development [16]. Fernandez and Robinson found student teachers were in favor of
microteaching sessions as it allowed them to practice pedagogical skills as one student teacher in their sample stated, "We spend a lot of time discussing theory. It's good to get a chance to try using some of it and get feedback on how well we did before we get our own classrooms." [84, p. 207]. It is possible microteaching might provide physics GTAs with opportunities to put pedagogy into practice while receiving feedback.

The microteaching model has been applied in some STEM GTA training programs and benefits have been documented in DBER literature. Acknowledging the importance of a coherent GTA training program on GTA teaching identity, Alicea-Muñoz, Sullivan, and Schatz designed a semester-long course for first-time physics GTAs [85]. Majority of the activities within the course take place before GTAs enter the classroom, including discussions of teaching expectations, pedagogy, and classroom management, and microteaching sessions related to different lab scenarios (referred to as lab simulation) and problem-solving sessions [85]. In their dissertation, Alicea-Muñoz found GTAs who participated in the course thought microteaching was a useful activity for preparing them to teach [17] similar to a finding in chemistry by Hume [86]. Additionally, Doucette, Clark and Singh reformed weekly preparation meetings for physics lab GTAs by including frequent microteaching sessions (referred to as "sabotage activity") in six out of ten weeks of training [18]. When comparing GTA teaching behaviors from the pre-training semester to the implementation semester, they found GTAs collectively increased the percentage of time spent engaged in open dialogue with and actively listening to students in the implementation semester [18]. In biology, Becker et al., designed a training program for biology recitation GTAs with eleven sessions scheduled throughout the semester [19]. The first five sessions
devoted the second hour of the session to practice of specific pedagogical skills like cold call (calling on a non-volunteer student by name to share an idea or respond to a question) while the remaining sessions required GTAs to work together to plan recitation lessons while integrating the pedagogical skills into that lesson [19]. They found GTAs implemented skills from the “drill” sessions into their teaching, but the frequency of GTAs’ use of the skills declined throughout the semester [19]. Cold call was one skill GTAs used in the classroom for the duration of the training (five weeks at the beginning of the semester), but their use of the skill declined the training was finished [19]. Becker et al. speculate deliberate practice is necessary to continue the use of pedagogical skills in the classroom. Thus, microteaching has a positive impact on a STEM GTA’s feeling of preparedness and teaching practices in the classroom, but further research is necessary.

However, microteaching sessions might not create authentic teaching experiences for STEM GTAs. He and Yan investigated the limitations of microteaching authenticity through interviews with pre-service teachers participating in their English as a Foreign Language education program [87]. They found pre-service teachers thought the different classroom environment with too many resources and their peers’ non-realistic acting as high schoolers were the most prevalent factors that contributed to the unauthenticity of their microteaching experience [87]. In their study about the role of technology in pre-service teacher preparation, Brown similarly reported some pre-service teachers find peers “playing” students is not as authentic of an experienced as they wished [88]. Brown continued their discussion by suggesting that technology can provide an alternative method to simulate more authentic classroom experiences [88].
Training Simulators Might Provide Ideal Environment for Skill Rehearsal

Training simulators can be used to scaffold learning within authentic situations and is supported by the situated cognition model presented in Chapter Two. By using simulated environments or situations for training, facilitators can provide trainees with multiple opportunities to practice the same situation without harming real people. For example, in the medical field, some resident programs include sessions with a virtual simulator paired with traditional hands-on practice to provide residents with multiple opportunities to practice complex surgical procedures without endangering real patients or having to wait for a real scenario to practice [89-91]. Simulator training may also provide better results for trainees than traditional training methods. In their meta-analysis on flight simulation used for training, Hays, Jacobs, Prince and Salas found student pilots performed the best on flight maneuvers when actual flight training was paired with simulator training [92]. Simulators can also support participants interest in engaging with content. In physics classrooms, PHET simulations might be used to encourage students to actively engage with physics content by exploring their own thinking about physics concepts through simulated situations [93-94]. Similarly, beginner teachers can benefit from simulator training by rehearsing complex pedagogical skills with avatar-students before trying the skills when teaching their real students.

Simulators like the mixed-reality integrated learning environment (MILE) and TLE TeachLivE™ (shortened to TeachLivE) were created for beginner teachers to gain real teaching experiences in a low-risk environment without impacting real students [21,26]. In this dissertation, physics GTAs participated in sessions with the TeachLivE simulator
to rehearse pedagogical skills. TeachLivE is considered a “mixed-reality” simulator because it has elements of virtual and physical reality integrated into the user’s experience [21]. To use the simulator, participants enter a physical room equipped with “real” classroom artifacts (i.e., board and markers) and a large monitor displaying a virtual classroom. When teaching a lesson, participants interact with the virtual class of five avatar-students in the same way they would talk with someone via video call.

Benefits of the simulator’s use for training on teacher skill development have been reported. In one study, a pre-service teacher who used the simulator to practice teaching a lesson about technology for elementary students was observed to improve their interactions with real students during class discussions [21]. As an exploratory study, Chini, Straub and Thomas implemented one session with the simulator into their physics Learning Assistants (LA) program and found LAs improved some strategies documented in their teaching self-reflections [25]. In parallel studies to this dissertation work, we reported an impact of incorporating multiple simulator sessions during one semester of training on STEM GTAs’ use of cold call and asking questions in their classrooms and labs [34,50]. Given such findings, we explore the use of the simulator as a tool for GTAs to practice using complex pedagogical skills such as asking questions (i.e., “Stretch-It”) and facilitating group work (e.g., managing group dynamics).

**Strategies To Support Instruction in Student-Centered Environments**

As suggested by Driessen, Knight, Smith, and Ballen, instructors who use student-centered instruction might lack fidelity in implementation because strategies that support student-centered instruction might not be well documented [53]. In this
dissertation, we provide GTAs with clear examples of pedagogical skills complementary to student-centered instruction. Here we discuss three pedagogical skills included in the teaching intervention: cold call, error framing, and Stretch-It. Each skill has the potential to create a safe learning environment for students to share their ideas.

In student-centered classes an instructor might consider using techniques that encourage a variety of students to share their ideas rather than just one student dominating the discussion. Cold call is a strategy that can be used in the classroom to increase the diversity of student voices heard [95]. In cold call, an instructor calls on a non-volunteering student to respond to a question or share an idea [96]. Here, the instructor can use cold call to set the classroom norm that student participation is expected and valued. Dallimore, Hertenstein, and Platt found the count of student volunteers increased throughout the semester when instructors frequently used cold call [39].

Since the mode of instruction depends on the contribution of student ideas, an instructor might consider using techniques that build on student ideas (similar to responsive teaching as described by [97]). Stretch-It is a questioning technique that extends the discussion beyond the student’s initial idea or answer [96]. This technique is also complementary to the learning objectives of inquiry-style labs or tutorials focused on sensemaking. For example, when exploring the use of physics tutorials, Koenig, Endorf and Braun found facilitators’ use of Socratic questioning instead of confirmation when checking in with students was linked to increased correct student responses to prompts about tutorial concepts [11]. Additionally, students who self-reported higher
learning gains also perceived their GTA to be interactive and to ask thoughtful questions [98].

Lastly, it is more than likely students will share an incorrect idea when an instructor uses cold call or Stretch-It, or even during group work. To encourage students to continue sharing their ideas despite sharing an incorrect idea, error framing can be used. Error framing is an apprehension reduction technique that can be used to re-direct student thoughts about having a correct answer to productively contribute to the classroom or group discussion [19,99]. In error framing, an instructor gives verbal statements that frame a student’s mistake as natural or beneficial to the learning process [46,99,100]. When investigating different training models, Bell and Kozlowski found use of error management training (incorporated error framing statements and emphasis on errors as positive to learning) contributed to higher levels of adaptive transfer (e.g., a student’s ability to apply knowledge to a new problem or setting) compared to error avoidance training [46].

**Student Feelings of Target Skills Need To Be Investigated**

While cold call can increase the voices heard in the classroom, it is possible that some students might not feel comfortable with the technique of cold call, as informed by fear of negative evaluation. England, Brigati, and Schussler found biology undergraduate students to self-report increased anxiety associated with their instructor’s use of cold call [101]. Similarly, Cooper, Downing, and Brownell investigated whether active learning practices increase or decrease a student’s anxiety; they found students only reported cold call to increase their anxiety [40]. Both studies were conducted for
lecture courses at research intensive universities. Downing et al. investigated the impact of active learning practices, including cold call, on student anxiety in biology community college classes and found students reported increased anxiety associated with their instructor’s use of cold call [47]. While community college courses might have smaller enrollment sizes compared to lecture courses at R1 institutions, there is limited research on the use of cold call in small enrollment physics classrooms. Since physics GTAs are expected to rehearse cold call in the simulator and might transfer that skill to their recitation and lab sections (informed by the positive impact of the simulator training on cold call use in the classroom [50]), it is necessary to investigate how students feel when their physics GTA uses cold call. Informed by the facework framework, error framing has the potential to mitigate negative student feelings towards the use of cold call (protective facework). For example, Downing et al. found students described their instructor’s use of error framing to reduce their fear of negative evaluation during whole class discussions [47]. In our previous study, we found students enrolled in the target courses (and the target course for chemistry, as part of the larger project) felt less anxious about answering a cold calling question when their chemistry or physics GTA paired cold call with another strategy [48]. We continue our investigation in this dissertation to investigate whether students perceived their physics GTA’s use of error framing to decrease their anxiousness in the classroom or lab.

The use of Stretch-It shifts the focus of who contributes knowledge in the classroom as instructors build on student ideas. While asking for reasoning, amongst other higher order cognition questions, might positively impact student learning outcomes (i.e., improved test scores [102]). Since there is limited research on how
students feel when their GTA uses questions in physics recitation and lab sessions, we investigate the impact of GTA questioning on student feelings. While we investigate the use of Stretch-It by GTAs in the simulator and in the classroom in Study One, we ask about how students felt about their GTA’s use of questioning in general in Study Three.

Chapter Summary

This chapter provided an overview of literature to support the investigation of using a mixed-reality simulator as a tool for physics GTA training. In summary, GTAs are crucial to STEM student learning and often teach in spaces where the expected instructional strategies may not mirror the GTAs’ own learning experiences. While GTAs benefit from receiving teacher professional development, the training they receive is limited. Namely, GTAs rarely receive opportunities to receive feedback about their teaching. Branching out into K-12 teacher preparation models, microteaching has been documented as a promising training activity to scaffold GTAs’ use of pedagogical skills in the classroom. To create a more authentic training experience to microteaching, a mixed-reality classroom simulator can be used to simulate realistic GTA-(avatar-)student responses. However, investigation of the impact of using a mixed-reality simulator for physics GTA preparation is limited. Thus, this dissertation adds to the existing research on physics GTA training models and applications of the mixed-reality classroom simulator, TLE TeachLivE™, by investigating the impact of rehearsal sessions with the simulator on physics GTAs and students.
CHAPTER FOUR: METHODOLOGY

While frameworks inform important processes of investigation like data collection and analysis and literature reviews point out the missing pieces of current research, research methodologies inform the study design (e.g., formulating research questions). In this chapter, we present two research methodologies that influenced the research study design, Action Research and Case Study Methodology. We also describe the context of the studies, including details about the teaching intervention and target courses. Finally, we include a positionality statement to account for the possible bias created by our positioning, namely our background and experiences.

Action Research

We used action research methodology to design this dissertation, as well as our larger project, as it fits our goals of improving the development and facilitation of ongoing GTA training in our home departments. Three main requirements of action research as informed by Carr and Kemmis are: 1) action research facilitators identify a social practice as subject to improvement, 2) the process of action research is a continuous cycle of planning, acting, observing, and reflecting, and 3) the project involves participants at every stage of the research [104]. An example of action research used in physics education research is the improvement of the ISLE curricula. In short, Etkina, Karelina, Murthy, and Ruibal-Villasenor used student artifacts (completed ISLE activities), classroom observations of students working on ISLE activities, and student lab report scores to investigate the alignment of the ISLE
curriculum with learning goals and to make changes to the use of ISLE in their own classrooms [104].

Our research team consisted of the course designers and former instructors of the target courses. While our team focused on both a chemistry laboratory and introductory physics mini-studio sequence, this dissertation focuses on the physics courses. Based on our experiences as course designers and instructors, we recognized that the curricula required physics GTAs to use skills that were not supported by existing GTA training, like asking questions without guiding students to the correct idea. Thus, we acknowledge the student-centered classroom to be a complex environment filled with various social dynamics (e.g., student-student interactions, TA-student interactions), and we identify the need to improve instructional support received by GTAs who teach in such environments. As training facilitators and action researchers, we worked together to improve the practice and comprehension of student-centered instruction while supporting social change in our departments [103,105].

In line with action research, we followed an iterative cycle for the design, implementation, and reflection of each training module. An example of our process for a chemistry training module focused on cold call and error framing is described in detail by Gereats et al. [34]. At multiple stages of the larger project, we elicited feedback from GTAs, in line with action research, to inform the direction of the study. For example, during the development of the training modules, we recruited experienced GTAs from the target courses to test the module and to provide us with feedback (i.e., their experience using the simulator and the usefulness of the target skill in the classroom).
We incorporated their feedback into the revised training module. In addition, the training and revision of the training was a cyclical process (another crucial piece as stated by [103]). Each simulator session followed the same cyclical process, described in a later section, and we continued to modify training sessions in response to GTA feedback during the actual training. We also personalized our feedback to GTAs and guided reflection. Also, facilitators were involved in multiple stages of data collection and analysis, as well as the dissemination of research to participants.

The type of action research used in this dissertation is considered technical action research, such that external members set goals for participants that might not be in alignment with participants’ goals [103]. Since we suggested to GTAs to use evidence-based teaching strategies during the simulator training, we set the goals for GTAs which might not align with what they think they need to work on. We recognize GTAs’ participation in the training might be affected by their involvement as technical participants. That is why we elicited GTA feedback and classroom observation data during each stage of training design, implementation, and reflection, similar to how Alicea-Muñoz collected survey data and classroom observation data when assessing their physics GTA training course [17].

**Case Study**

In Study Two and Three, we applied a case study methodological approach to investigate details of GTA-student interactions that are not described by the classroom observation tool (described later in this section), including GTA awareness/resolution of student group challenges and student feelings about their GTA’s use of the target skills.
Case study methodology can be used when a study requires an in-depth exploration of a group or phenomenon [105].

In detail, case study methodology can be used when investigating real-world scenarios and situations when the context of the study and the investigated phenomenon (in this case GTA-student interactions) cannot be disentangled [106]. In this dissertation, the implemented teaching intervention occurred alongside other training received by GTAs and while GTAs were teaching. Thus, students were interacting with GTAs when GTAs were forming their teaching identity. Additionally, GTA-student interactions occurred in the classroom or lab where the curriculum, physical space, and student interactions with their peers might influence student perception of GTA-student interactions (i.e., as described by the instructional capacity framework [107]). Since GTA-student interactions occurred in the real-world, we applied the case study methodology when investigating the student perception of group challenges and their GTA’s group management strategies in Study Two. We also implemented case study methodology in Study Three when exploring student feelings of their GTA’s use of target pedagogical skills from the simulator training in their class.

Roman and Uttamchandani have argued for qualitative approaches to be used when investigating instructor-student interactions in smaller interactive courses [108]. While case study research methodology is not limited to use with qualitative data [106], we use a qualitative case study approach in Study Two and Three to explore the depth of GTA-student interactions in mini-sections. Thus, for Study Two and Study Three we follow a qualitative case study approach. This approach informed us about the type of
data to collect and analyze when obtaining the richness required for investigating student perceptions of group work and the target skills. By exploring a small set of student interviews, we can answer the how and why questions associated with student experiences of group work and their GTA’s use of the target skill. An example of a qualitative case study approach is a study exploring physics tutorial GTA behaviors. Goertzen, Scherr, and Elby used case study methodology to investigate snippets of video-recorded classroom observations and GTA interviews for fine-grained details of GTA behaviors in the classroom and the cognitive decision behind that GTA behavior [109].

Furthermore, we used a multiple-case study design (also referred to as comparative case study) such that we compare findings across individuals in our sample and draw conclusions across all cases [106]. In Study Two and Three, students might respond in different ways, and we report the common experiences shared by individuals. Therefore, we present individual cases as support for each claim and for pieces of discussion.

**Positionality**

This dissertation investigates unique pieces from a larger project investigating the effectiveness of implementing rehearsal sessions with a mixed-reality simulator into STEM GTA training. Throughout the timeline of collecting and analyzing data for this dissertation, I (C.M.D.) have held several roles including graduate student peer, training facilitator and researcher. The training facilitator role is not limited to leading rehearsal
sessions as I have also collaborated with a post-doc (T.W.) to facilitate weekly preparation meetings for the target courses.

As a training facilitator for the teaching intervention and leader of weekly preparation meetings, I recognize the power dynamic between the GTAs and myself. It is possible observing and providing feedback to the GTAs during their rehearsal sessions might have resulted in GTAs perceiving me as an authoritative figure. A similar result might have occurred during weekly preparation meetings. GTAs might have felt nervous while practicing in the simulator and receiving feedback from me because they do not want to be negatively evaluated by their trainer. To help mitigate the potential negative impact of GTAs feeling like they are being evaluated, I made statements like, “The simulator is a space for you to practice” and “We care about you and your role as a GTA and therefore we are here to support you while you are teaching.”

As a researcher, I conducted classroom observations and interviews. During classroom observations, GTAs might have felt more inclined to use the target skills because of my presence in the classroom. Similar to receiving feedback, they might have felt nervous to “perform well” in the classroom since they do not want to be seen as not teaching according to how the department expects. To help mitigate the negative impact of my presence during classroom observations, I talked with GTAs before and after class to build rapport. I also made statements like, “We are just curious if you find the teaching skills useful in the classroom” or “We are just curious if the training is supporting you in your role as a GTA and if not, we can make changes to best support you.” While conducting interviews for either GTAs or students, participants might feel
bad for not agreeing with us about the usefulness of the target skills or expressed dissatisfaction with their GTA or mini-studio section. To break down the power I had as a researcher I ensured participants that their ideas and experiences are valid and useful contributions to our research. I also ensured to participants that their responses to our interview questions are anonymous and that I am interested in their experiences.

My role as a graduate student peer might have softened GTAs’ perception of me as an authoritative figure, thus challenging the power hierarchy. It is possible GTAs might have felt more comfortable talking with me about the training or their experiences in the classroom because I am their peer. However, some GTAs might have only viewed me as a peer and would take feedback more seriously if a faculty member supplied it. As a graduate student, I might have also seemed less intimidating than a faculty member to students when conducting interviews.

**Teaching Intervention: GTA Simulator Training**

During Fall 2019 and Spring 2020, we integrated rehearsal sessions with the simulator into physics GTA training for those teaching the target courses (described below). GTAs teaching in either semester had a variety of experience teaching the target courses and other physics courses.

In Fall 2019, GTAs participated in four sessions with the simulator throughout the semester. The first three sessions focused on practicing different pedagogical skills while the last session focused on integrating target skills from previous sessions into one coherent session. Each session was held in-person.
In Spring 2020, there were a mixture of returning GTA from Fall 2019 and GTAs new to the training iteration. Returning GTAs continued their simulator training from Fall 2019 by participating in three sessions focused on practicing all skills together. New GTAs followed Fall 2019’s simulator training model. The first three session were held in-person while the last session was held remotely due to the COVID-19 pandemic.

Before each session, we provided GTAs with multiple documents describing what to expect when interacting with the simulator, details about the target pedagogical skills for the session, and the lesson avatar-students are expected to work on for the session. An example of each document is provided in Appendix A.

During each session, GTAs enter a physical room to interact with the virtual classroom that is displayed on a screen or monitor in the same room, as presented in Figure. 3. GTAs could interact with physical classroom artifacts around them such as markers, a whiteboard, and student work on paper.

Figure 3: Picture of a researcher (T. W.) interacting with the simulator. Left Image: Virtual classroom projected on a screen in front of simulator user; Right Image: Researcher (T. W.) interacting with simulator in a physical room.
To create a realistic experience with the simulator, we included common student ideas and reasoning into avatar-student dialogue, and we leveraged the personalities of the five avatar-students (Kevin, CJ, Ed, Maria, and Sean) to create dysfunction within student groups (for the third session, focused on group facilitation). Additionally, we framed each simulator session to be a snapshot of teaching in either recitation or lab such that avatar-students worked in groups on activities similar to the mini-studio curricula. In the cold call and error framing session (session one), we used the Physics I tutorial lesson, “Timmy Down the Well” which focused on students reasoning through Newton’s laws of motion and common sense. During the session focused on Stretch-It (session two), we used the pendulum lab from the “Thinking Critically” lab curriculum [110]. For the group management strategies session (session three), we used a tutorial from the Physics II mini-studio that included students building three different direct current (DC) circuits and comparing the brightness of the lightbulbs in each circuit. Additionally, in the final simulator session, we used the same tutorial activity from the previous simulator session (session three).

Figure 4: Schematic of the Cyclical Procedure for a Simulator Session
Simulator sessions followed a cyclical procedure of deliberate practice: practice teaching in simulator (Round 1 and 2) with facilitator-led feedback and reflection in between rounds, as depicted in Figure 4. GTAs practiced in the simulator for approximately seven minutes each round.

During feedback and reflection, facilitators posed questions about the GTAs’ skill use in the simulator (i.e., “Do you feel like you were able to use the target skill?”) and their use of the skill in the classroom (i.e., “How would you use this in your class?”). We provided feedback to GTAs including opportunities where they could have used the skill during their practice before the next round. GTAs typically participated in groups of three such that each GTA individually rehearsed in the simulator while the other GTAs in their group watched them interact with the simulator. We have provided one round of dialogue for one GTA in Appendix A as an example of how GTAs might have interacted with the avatar-students.

**Context**

*Description of Target Courses*

Target courses for this dissertation are introductory algebra-based physics “mini-studio” sections. Mini-studio refers to a combined tutorial and lab session lead by physics GTAs [3,111]; the mini-studios are part of a two-semester sequence (Physics I and Physics II) [112]. During each mini-studio session, students [113] are expected to work in small groups (3 – 4 students) on tutorial worksheets adapted from the “Tutorials
on Physics Sense-making” [7], a quiz, and an inquiry-style lab report based on the ISLE curriculum [8].

Before Fall 2019, GTAs taught tutorial and lab in the same room. In Fall 2019, the Physics I tutorial was moved to a different room while the lab remained in the original classroom. This same change occurred for Physics II mini-studio sessions in Spring 2020. While we do not investigate the impact of changing the physical classroom on the interactions between GTAs and their students in this dissertation, we reported the impact of the classroom on GTA and student behaviors in a previous study [114]. Thus, we recognize that the physical classroom might influence our studies and is a limitation for this dissertation.

*Pre-Existing GTA Training*

Graduate students new to the teaching assistant program attended a one-semester pedagogy seminar during their first semester as a GTA, instructed by an author of this paper (J.J.C.). Each week GTAs read articles about general education topics (e.g., Bloom’s taxonomy, student mindset about learning, etc.), wrote teaching reflections and participated in face-to-face Socratic discussions with other new GTAs.

In addition, all GTAs are required to attend weekly preparation meetings facilitated by a head GTA or training facilitator. Weekly preparation meetings for Physics I and II mini-studio GTAs were facilitated by two authors of this paper (T.W. and C.M.D.) during Fall 2018 and the first author (C.M.D.) during Fall 2019. Meetings were run the same way both semesters. During each meeting, we encouraged GTAs to work together as a whole group or in small groups to discuss the tutorial worksheet and lab for the following
week. While working in groups, GTAs shared their experiences in the classroom and common student responses.

Data Collection and Analysis

In this dissertation, we collected and analyzed simulator rehearsal sessions, classroom observation data, and undergraduate student interviews. We use data across three different semesters when investigating the over-arching research question. In Study One, we analyze simulator rehearsal sessions from Fall 2019 and classroom observations from Fall 2018 and Fall 2019. In Study Two and Study Three, we analyze student interviews from Spring 2020. A timeline for data collection from cohorts participating in Fall 2018, Fall 2019, and Spring 2020 is presented in Figure 5.

Figure 5: Timeline of all semesters. A – Pre-simulator Semester; B – Simulator Training Semester 1; C – Simulator Training Semester 2
Classroom Observation Protocol

In Study One, we used a modified version of the Laboratory Observation Protocol for Undergraduate STEM (LOPUS) to document GTA and student behaviors in the tutorial and lab portions of the mini-studio. Velasco et al. created LOPUS as a tool to describe learning and instruction in postsecondary chemistry labs [115]. While Velasco et al. used LOPUS when observing traditional-style (confirmation) chemistry labs, we chose to use LOPUS because the GTA and student behaviors matched our expectations of those in the mini-studio.

Our modified version of LOPUS has 31 codes described in detail in our previous work [116]. In Study One, we only focus on the use of four codes when analyzing classroom observation data. Those codes are TA Posing Question (PQ), One-on-One TA Posing Question (1o1_TPQ), Stretch-It: Explain Logic (Explain Logic), and Stretch-It: Follow-Up (Follow-Up). PQ occurred was coded when a GTA asked a question in front of the whole class. 1o1_TPQ was coded when a GTA asked a question in a small group or individual student setting. Explain Logic or Follow-Up was coded along with either PQ or 1o1_TPQ, depending on the setting the Stretch-It question was asked.

Using the Observation Protocol

We conducted live classroom observations using the modified LOPUS. During each observation, a researcher coded if a GTA or student behavior occurred in a two-minute interval and continued coding for the duration of the mini-studio section. To limit contact between researchers, GTAs, and students, the GTA wore a lapel
microphone connected to a transmitter. The researcher listened to conversations between GTA and students through a receiver. We discuss the limitations of using this technology in Study One. Codes are not mutually exclusive. For example, a GTA might actively observe a student group, initiate an interaction with the group, and pose a question to the group, all within the same two-minute interval.

**Pedagogical Skills Added To Protocol**

We modified the protocol to include codes for the target pedagogical skills from the simulator training. Since LOPUS differentiates between level of GTA-student interactions, one-on-one (individual or small groups of students) and whole class, we created two codes for cold call since a GTA might use cold call in either setting. We created two codes for Stretch-It to differentiate between the two main categories: Explain Logic and Follow-up. However, we did not create a Stretch-It code specific to instructional setting. When coding Stretch-It we also coded that the GTA posed a question which already is split into two codes for each setting. We created only one code for error framing. Group management strategies were not documented using LOPUS (details in Study Two).

**Analyzing Classroom Observation Data**

After each classroom observation, an observation log was created. The log was a table that reported whether or not any of the 31 codes were coded during any two-minute interval for the duration of the observation. Each column represented a code while each row represented one two-minute interval. For example, if the duration of the
observation was 60 minutes, the log would display a table with 31 columns and 30 rows. If a code was coded during a two-minute interval a “1” appeared in the cell for the code and the corresponding row. If a code was not coded during a two-minute interval a “0” appeared in the cell for the code and the corresponding row. By summing each column, we obtain the total number of two-minute intervals a code occurred for that observation. In Study One, we obtained the total number of two-minute intervals EL and Fol occurred for each observation by summing the column for each code.

Analyzing Student Interview Data

In this dissertation, two different methods of analysis are used to analyze student interview data. In Study Two, a list of a priori codes, namely the scenario prompts, we used when analyzing the student interviews for instances where they experienced a group challenge descriptive of those prompts [117]. In Study Two and Study Three, thematic analysis was used to allow the themes to emerge from the student interviews about their experiences with alternative group challenges, feelings about the target skills, and ways GTAs responded to their incorrect ideas [118].

Chapter Summary

This chapter described research methodologies used in this dissertation. In Study One, an Action Research methodology was used when creating and implementing new professional development for physics GTAs. In Study Two and Three, case study methodology was used when investigating details about GTA-student
interactions in mini-studio sections not described by the observation protocol. The observation protocol used in Study One is also described here.

In addition, the simulator training intervention was described in detail for context for professional development received by focused GTAs in all three studies. A description of the target courses is also provided for context about the learning environment in all three studies.
CHAPTER FIVE: RESEARCH QUESTIONS

This dissertation aims to investigate the impact of using a mixed-reality classroom simulator for GTA training on GTAs and students. To assess the impact of simulator training on GTAs, the framework detailed by Reeves et al. suggests investigating GTA teaching behaviors in the classroom/lab because of the direct impact of GTA outcomes on students. The impact of GTA teaching behaviors on students motivates the investigation of student responses to their GTA's behaviors in the classroom/lab, including the use of the target skills. Therefore, informed by Reeves et al., this dissertation focuses on answering the following over-arching question: What is the impact of teaching rehearsal in a mixed-reality classroom simulator on physics GTAs and their students in active learning lab and recitation sessions?

The over-arching research question will be addressed by presenting three unique studies. Study One will focus on the impact of the simulator training on GTAs while Study Two and Study Three focus on exploring the impact of the simulator training on students.

To assess the impact of simulator training on GTAs' use of the target skill, Stretch-It, the framework detailed by Reeves et al. suggests investigating GTAs' use of Stretch-It in their classroom/lab. Since GTA responsiveness regulates the impact of training design on GTA teaching outcomes, it is equally important to investigate how GTAs respond to the teaching intervention [27]. Therefore, the following questions will be addressed in Study One:
1. Do GTAs change their use of questioning in simulator sessions after rehearsing the Stretch-It skill?

2. What is the impact of practice during simulator sessions on GTAs’ use of questioning in their actual classroom?

Simulator rehearsal video data and classroom observation data were analyzed using a priori coding methods to investigate GTAs’ use of Stretch-it during simulator sessions and in the classroom.

The framework outlined by Reeves et al. describes the impact of GTA teaching practices in the classroom on student outcomes. Since GTAs are expected to rehearse interacting with avatar-student groups with group dysfunctions, students might perceive their GTA to help resolve their group dynamics challenges in their mini-studio section. In group work, students might not feel comfortable sharing an idea because students are afraid of being negatively judged for having a wrong idea, as described by fear of negative evaluation [38, 40]. GTAs can take actions to mitigate the impact of negative student experiences and feelings on student self-perceptions during group work, especially those created by fear of negative evaluation by doing facework [41,42]. Study Two aims to explore the student perception of group dynamics challenges while investigating how students perceive their GTA to help resolve their group issues by asking the following questions:

1. What group challenges do students experience in the physics mini-studio sections?
2. How do students perceive their GTA to recognize and help resolve their group challenge?

3. How do students perceive their GTA’s role in the classroom and how is it related to their GTA resolving their group challenges?

Student interview data will be analyzed using a codebook comprised of a priori codes (scenario prompts and mentor/instructor) and emergent themes. Emergent themes will be found using thematic analysis to allow alternative group challenges to emerge from the student interview data.

Also informed by Reeves et al., Study Three aims to investigate the impact of the simulator training on students. Like Study Two, Study Three uses fear of negative evaluation and facework to motivate the investigation of how students feel when their GTA uses or might use the target skills in their mini-studio section. Thus, Study Three aims to answer the following questions:

1. Do students perceive their GTA to use the target skills from the simulator training in their classroom or lab?

2. How do students feel when their GTA used the target skills in the classroom or lab?

3. What other ways do students report their GTA to respond to incorrect ideas, and how do they feel about them?

Student interview data will be analyzed using a codebook comprised of a priori codes (cold call and error framing) and emergent themes. Emergent themes will be
found using thematic analysis to allow student feelings and other ways for GTAs to respond to an incorrect student idea to emerge from the student interview data.
CHAPTER SIX: STUDY ONE – IMPACT OF HIGH-INTENSITY TRAINING WITH A MIXED-REALITY SIMULATOR ON GTAS’ USE OF QUESTIONING

Introduction

University physics departments often depend on GTAs to teach the recitation and laboratory sections of introductory physics courses. Responding to national calls to transform instruction to support student learning [1], many physics departments have adopted student-centered curricula in recitations and labs. Such curricula require complex pedagogical skills, some of which GTAs are unfamiliar or uncomfortable implementing. Furthermore, STEM GTA training varies across universities and across departments within the same university [35,65,70]. We created opportunities for GTAs to practice complex pedagogical skills, like asking questions, using a mixed-reality classroom simulator. Questioning can shift the focus of who supplies the knowledge in the classroom from the instructor to the students, promoting student engagement [95] and supporting student learning [11,102]. In this study, we investigate GTAs’ use of questioning during mixed-reality simulator sessions focused on questioning, subsequent mixed-reality simulator sessions with other foci, and in their actual classroom. We introduced GTAs to a particular questioning skill called “Stretch-It” [96], which provides ideas to formulate questions that encourage students to explain their logic or follow up on their understanding of the focal skill, as discussed in more detail below.
Frameworks Utilized

As mentioned in the Frameworks chapter, we utilized the GTA Professional Development Evaluation and Design framework to guide the design and evaluation of our GTA training program. This framework conceptualizes relationships between GTA training outcomes that impact students and factors that influence GTA outcomes [27]. Namely, the framework informs us that GTA teaching practices used in the classroom influence student outcomes, and that training design, implementation of training, and GTA characteristics impact GTA teaching practices. Thus, we report GTA engagement during training sessions and GTA use of the target skill (Stretch-It) in the classroom in the findings section. We also allude to GTA characteristics influence on our findings in the discussion section.

Additionally, we used situated cognition to inform the activities included in the training that would impact GTAs’ use of pedagogical skills in the classroom [20]; We used a simulator to simulate a virtual class for GTAs to interact with when rehearsing pedagogical skills.

Research Questions

Since STEM GTA teaching practices impact students, we explored the use of a simulator to help GTAs develop pedagogical skills like useful questioning strategies such as Stretch-It. However, literature measuring the impact of using simulation for STEM GTA training is limited. Thus, we investigated GTAs use of Stretch-It during simulator sessions and in their classroom. We ask the following research questions:

1. Do simulator sessions support to rehearse Stretch-It questioning?
2. What is the impact of practice during simulator sessions on GTAs’ use of questioning in their actual classroom?

Methods

Developing GTA Training through Action Research

In line with action research [103,105], our research team consisted of the course designers and former instructors of the target courses. Here, we focus on GTAs teaching physics. Based on our experiences as course designers and instructors, we recognized that the curricula required physics GTAs to use skills that were not supported by existing GTA training. We followed an iterative cycle with the design, implementation, and reflection informed by GTA feedback of each training module. An example of our process for a chemistry training module focused on cold call and error framing is described in detail by Gereats et al. [34].

Positionality

We recognize the variation of academic positions within our research group that contribute to different power dynamics between ourselves and our participants. C.M.D. executed several roles in the project and target courses: graduate student peer, training facilitator and classroom observer. In addition, C.M.D. and T.W. led weekly preparation meetings for target courses. J.J.C. taught the pedagogy course, which most of the participating GTAs had either previously or were currently taking. A.A.G., C.N. and E.K.H.S. occupied similar roles in the chemistry department but were lesser known by and had less power in relation to the physics GTAs.
Target Courses – Recitation and Lab Sections

The target courses are introductory algebra-based physics “mini-studio” (combined recitation and lab) sections [3,111] which are part of a two-semester sequence (Physics I and Physics II) [98]. During each mini-studio session, students [99] are expected to work in small groups (3 – 4 students) on tutorial worksheets adapted from the “Tutorials on Physics Sense-making” [7], a quiz, and an inquiry-style lab report based on the ISLE curriculum [8].

Participants

Pre-Simulator Training Semester

Before implementing the simulator training intervention, we observed eight Physics I or II mini-studio GTAs teaching in their real classrooms during Fall 2018. The GTAs varied in their experience teaching mini-studio (reported in Table 1), as well as other teaching experience, gender, and nationality. Fall 2018 GTAs consented to observation log and audio-recordings of their observations.

Table 1: Pre-Simulator Training Semester Participants

<table>
<thead>
<tr>
<th>GTA</th>
<th>Mini-Studio Course</th>
<th>GTA Prior Teaching Experience</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Physics I</td>
<td>New</td>
</tr>
<tr>
<td>II</td>
<td>Physics I</td>
<td>Mini-Studio</td>
</tr>
<tr>
<td>III</td>
<td>Physics I</td>
<td>New</td>
</tr>
<tr>
<td>IV</td>
<td>Physics I</td>
<td>New</td>
</tr>
<tr>
<td>V</td>
<td>Physics II</td>
<td>New</td>
</tr>
<tr>
<td>VI</td>
<td>Physics II</td>
<td>Mini-Studio</td>
</tr>
<tr>
<td>VII</td>
<td>Physics II</td>
<td>New</td>
</tr>
<tr>
<td>VIII</td>
<td>Physics II</td>
<td>New</td>
</tr>
</tbody>
</table>
Table 2: High-intensity Simulator Training Semester Participants

<table>
<thead>
<tr>
<th>GTA</th>
<th>Mini-Studio Course</th>
<th>GTA Prior Teaching Experience</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Physics I</td>
<td>New</td>
</tr>
<tr>
<td>2</td>
<td>Physics I</td>
<td>New</td>
</tr>
<tr>
<td>3*</td>
<td>Physics I</td>
<td>New</td>
</tr>
<tr>
<td>4°</td>
<td>Physics I</td>
<td>New</td>
</tr>
<tr>
<td>5°</td>
<td>Physics I</td>
<td>New</td>
</tr>
<tr>
<td>6</td>
<td>Physics I</td>
<td>New, Other Courses</td>
</tr>
<tr>
<td>7</td>
<td>Physics I</td>
<td>New, Other Courses</td>
</tr>
<tr>
<td>8†</td>
<td>Physics I</td>
<td>Mini-Studio</td>
</tr>
<tr>
<td>9°</td>
<td>Physics I</td>
<td>Mini-Studio</td>
</tr>
<tr>
<td>10</td>
<td>Physics II</td>
<td>New</td>
</tr>
<tr>
<td>11*</td>
<td>Physics II</td>
<td>New</td>
</tr>
<tr>
<td>12</td>
<td>Physics II</td>
<td>Mini-Studio</td>
</tr>
<tr>
<td>13</td>
<td>Physics II</td>
<td>Mini-Studio; Other Courses</td>
</tr>
<tr>
<td>14</td>
<td>Physics II</td>
<td>Mini-Studio; Other Courses</td>
</tr>
</tbody>
</table>

*Partial Data. °Classroom Observation Data Only. †Simulator Session Data Only. GTA 4, 6, and 9 only had a log for classroom observations.

High-Intensity Simulator Training Semester

During the high-intensity simulator training semester in Fall 2019, we facilitated training for and observed sixteen GTAs who taught either Physic I or II mini-studio sections (there was no overlap between the Fall 2018 and Fall 2019 cohorts). Thirteen GTAs (81.3% participation) consented to their classroom observation data and eleven GTAs (68.8% participation) consented to both simulator and classroom observation data to be used for research purposes. High-intensity training participants varied in their experience teaching mini-studio (displayed in Table 2), as well as other teaching experience, gender, and nationality.

Some GTAs have taught other courses at this university before teaching mini-studio sections, like traditional physics labs (i.e., confirmation labs) or studio-mode physics courses with a faculty instructor. As informed by Reeves et al., previous
teaching experience might influence GTA buy-in to the additional professional development [27].

**Pre-existing GTA Training**

Graduate students new to the teaching assistant program attended a one-semester pedagogy seminar during their first semester as a GTA, instructed by an author of this paper (J.J.C.). Each week GTAs read articles about general education topics (e.g., Bloom’s taxonomy, student mindset about learning, etc.), wrote teaching reflections and participated in face-to-face Socratic discussions with other new GTAs.

In addition, all GTAs are required to attend weekly preparation meetings facilitated by a head GTA or training facilitator. Weekly preparation meetings for Physics I and II mini-studio GTAs were facilitated by two authors of this paper (T.W. and C.M.D.) during Fall 2018 and the first author (C.M.D.) during Fall 2019. Meetings were run the same way both semesters. During each meeting, we encouraged GTAs to work together as a whole group or in small groups to discuss the tutorial worksheet and lab for the following week. While working in groups, GTAs shared their experiences in the classroom and common student responses.

In our study, we used a mixed-reality classroom simulator to create an environment for practicing questioning skills. We operationalized the term questioning skills to be posing Stretch-It questions to (avatar-)students. Stretch-It has two main categories with two sub-categories for each main category: Explain Logic (“Ask for Work” or “Ask for Evidence”) and Follow-Up (“Ask about Analogous Situation” and “Ask for Another Way to Answer”). All categories of Stretch-It are displayed in Table 3.
### Table 3: Questioning Skills Presented in Simulator Intervention

<table>
<thead>
<tr>
<th>Target Skill</th>
<th>Skill Description</th>
<th>GTA Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stretch-It:</td>
<td>Ask students to explain the reasoning behind their answer or idea, after asking</td>
<td>GTA: <em>How did you measure the period?</em></td>
</tr>
<tr>
<td>Explain Logic [91]</td>
<td>a question or the student shares an idea</td>
<td></td>
</tr>
<tr>
<td><strong>1. Ask for Work</strong></td>
<td>Ask students to explain how they got the answer</td>
<td>GTA: <em>Okay. So, CJ and Ed, do you think your data are reliable, more</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>valid than them [other group]? [CJ: Yeah.] Uh, how do you determine</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>your data is more reliable?</em></td>
</tr>
<tr>
<td><strong>2. Ask for Evidence</strong></td>
<td>Ask students why they think their answer is valid or not valid</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stretch-It:</td>
<td>Ask students to stretch the boundaries of their knowledge and check for</td>
<td>GTA: <em>So, think about like a dart board, if you have a really accurate</em></td>
</tr>
<tr>
<td>Follow-Up [91]</td>
<td>integration, after asking a question or the student shares an idea</td>
<td><em>collection of darts, you set, like, they’ve all like gotten really close to</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>the center. Right?</em>[Sean: Uh-uh. Right.] So, what do we call it when all the*</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>darts have hit the same spot, but it’s not the center?</em></td>
</tr>
<tr>
<td><strong>1. Ask about</strong></td>
<td>Ask students to relate the same concept or skill in a different context</td>
<td>GTA: <em>That sounds very textbook. Could you say that in um slightly</em></td>
</tr>
<tr>
<td>Analogous Situation</td>
<td></td>
<td><em>layman’s terms or without so much technical jargon? Or could you explain</em></td>
</tr>
<tr>
<td><strong>2. Ask for</strong></td>
<td>Ask students to explain or answer in a different way</td>
<td><em>that differently?</em></td>
</tr>
<tr>
<td>Another Way to Answer</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The simulator training model is discussed in detail in the Methodology Chapter of this dissertation. In short, GTAs entered a physical room with classroom artifacts (e.g., board, markers, etc.) and interacted with a virtual class of five avatar-students.
displayed on a screen (a visual representation is presented in Figure 3 in the Methodology Chapter). During each rehearsal session up to three GTAs entered the simulator room and took turns leading a discussion about the lesson with the avatar-student class while practicing the target skills. After each GTA finished their first rehearsal (Round 1), facilitators provided feedback and asked reflection questions about the target skill. Then, the simulator was reset and GTAs participated in a second round of rehearsal. Before GTAs left for the day, facilitators provided another round of feedback and asked questions pertaining to the target skill use in the classroom. We have provided one round of dialogue for one GTA in Appendix A as an example of how GTAs might have interacted with the avatar-students. In the questioning skills session, we implemented an activity from the “Thinking Critically” lab curriculum [110] because we intended to recruit GTAs who teach with a different curriculum. Additionally, in the final simulator session, we used the same recitation activity from the previous simulator session (session three).

Data Collection and Analysis

Classroom Observations

Before implementing simulator sessions into GTA training, we observed eight mini-studio GTAs teaching in their classroom during five lessons in Fall 2018 (total of 32 observations) to tease out the influence of the course (e.g., curriculum) on GTAs’ behaviors when investigating the impact of the simulator training on GTAs’ classroom practices. In Fall 2019, during the high-intensity simulator training, we observed thirteen mini-studio GTAs teaching in their classroom (total of 37 observations) during three
lessons. Fall 2018 and Fall 2019 cohorts are distinct such that there is no overlap in participation. All classroom observations in Fall 2018 were audio-recorded. Fall 2019 GTAs had the option to opt out of audio-recordings; three Fall 2019 GTAs consented to only observation logs to be used for research (indicated in Table 2). Both timelines are available in Figure 6.

For our larger project we (C.M.D., A.A.G, T.W., and C.N.N.) used a modified version of the Laboratory Observation Protocol for Undergraduate STEM (LOPUS) to capture GTA and student behaviors in the classroom [101]; we have previously discussed our modifications to LOPUS and clusters of GTA and student behaviors observed in the target courses [116]. During Fall 2019, we added Stretch-It categories to our observation protocol. We collapsed the four Stretch-It sub-categories into two main categories, Explain Logic ("Ask for Work" and "Ask for Evidence") and Follow-Up ("Ask about Analogous Situation" and "Ask for Another Way") to reduce cognitive load on observers. We focus on those codes here.

In addition, we explored inter-rater reliability (IRR) by calculating Gwet’s AC1 values for pairs of coders; Gwet’s AC1 is a metric that can be used to interpret IRR and is robust for data with low-trait prevalence [119,120]. Averaged Gwet’s AC1 values across pairs of coders for Explain Logic was 0.94 before discussion and 1 after discussion, and averaged Gwet’s values across coders for Follow-Up was 0.99 before discussion and 0.99 after discussion. We also calculated Gwet’s AC1 values for two GTA behaviors mentioned later in this section, PQ and 1o1_TPQ. Averaged Gwet’s AC1 values for PQ and 1o1_TPQ were 0.74 and 0.95, respectively. Values between
0.61 and 0.8 can be considered as moderate agreement while values above 0.81 can be interpreted as near perfect agreement [120].

Since we did not originally code for Stretch-It in Fall 2018, we revisited the Fall 2018 classroom observation data. To code Fall 2018 observation data for Stretch-It, we matched time codes on audio files with the corresponding observation logs (two-minute intervals) for each GTAs’ observation. For each two-minute interval PQ or 1o1_TPQ was coded as indicated on each observation log, we listened to the two-minute interval on the corresponding audio file for Explain Logic and Follow-Up Questions. Fall 2018 classroom observation data was coded for Stretch-It after investigating IRR for Fall 2019 classroom observations and Fall 2019 simulator sessions.

Figure 6: Timeline for Both Cohorts. A) (Top) Timeline of Fall 2018 Pre-Simulator Intervention. B) Timeline of Fall 2019 High-intensity Simulator Intervention
Preparing Classroom Observation Data for Comparison

To assess the impact of the simulator training on GTAs’ implementation of Stretch-It questions in the classroom, we compared high-intensity simulator training (Fall 2019) classroom observation data to pre-simulator training (Fall 2018) classroom observation data for each category of Stretch-It [121]. For each observed lesson we added the number of two-minute intervals a Explain Logic or Follow-Up question was coded for all GTAs observed for that lesson, then we divided by the number of GTAs observed for that lesson. We calculated this value for each lesson observed in Fall 2018 and 2019.

Sessions with the Mixed-Reality Classroom Simulator

Simulator sessions were audio and video-recorded and transcribed by the first author (C.M.D.). Three authors (C.M.D., A.A.G and E.K.H.S.) co-coded the transcripts in two stages using a list of a priori codes, as displayed in Table 3. First, we coded segments of the transcripts (~1%) for training purposes. Here, each researcher independently found examples of Stretch-It questions which were discussed until agreement was reached [122]. During the second stage of coding, two researchers (A.A.G. and E.K.H.S.) coded two different sets of transcripts (~8% of the transcripts each) while the first author (C.M.D.) independently coded both sets (~16% of the transcripts). The grain-size for coding transcripts was GTA turn-of-speech (when a GTA was talking with avatar-students). Similar to classroom observations, we calculated Gwet’s AC1 for pairs of coders (A.A.G., E.K.H. S., and C.M.D.) to investigate IRR. Gwet’s AC1 values for pairs of coders before discussion ranged between 0.94 and 1.
and 0.96 to 1 after discussion, where values larger than 0.81 can be interpreted as nearly perfect agreement [120].

Findings

GTAs’ Use of Stretch-It during High-intensity Simulator Training

Stretch-It was the focal skill of the second simulator session in the high-intensity training semester. GTAs’ use of Stretch-It in Round 1 and 2 of this simulator session are shown in Figure 7, which uses a stacked bar chart to display both the total Stretch-It count per category across GTAs as well as each GTA’s contribution to the total. Black boxes represent Physics I GTAs and gray boxes represent Physics II GTAs [123].

GTAs Used Different Types Of Stretch-It in Different Rounds

In Round 1, nine out of eleven GTAs asked a Stretch-It question. We found “Ask for Work” to be the most common type of Stretch-It question posed across GTAs (seven out of eleven GTAs) followed by “Ask for Evidence” (four GTAs) and “Ask for Another Way to Answer” (two GTAs). We did not observe GTAs to use “Ask about an Analogous Situation”.

After facilitator-led feedback and reflection, we observed all eleven GTAs to use Stretch-It during their second round. We observed all four types of Stretch-It questions to be asked by at least two GTAs. Eight GTAs used both “Ask for Evidence” and “Ask for Work” questions; “Ask for Evidence” was used more frequently than “Ask for Work”, a reversal compared to Round 1. We also observed three GTAs to use “Ask about an Analogous Situation” during Round 2 (no GTAs used this question-type in Round 1).
Figure 7: GTAs’ Use of Stretch-It during Questioning Skills Session: (A) (Left): Round 1; (B) (Right): Round 2. W – “Ask for Work”, E – “Ask for Evidence”, AS – “Ask about Analogous Situation”, AW – “Ask for Another Way to Answer”

The total Stretch-It questions asked across GTAs increased from 22 to 33 turns-of-speech across rounds. Additionally, we found six GTAs (1, 6, 7, 10, 11, and 12) increased the number of turns-of-speech in which they asked Stretch-It questions from Round 1 to Round 2 while two GTAs (3, 8) asked Stretch-It questions in the same number of turns-of-speech in both rounds. However, three GTAs (2, 9 and 13) decreased the number of turns-of-speech in which they asked a Stretch-It question across rounds. We note that not only the quantity but also the quality of questions can be used to evaluate the training effectiveness. Quality of questions require in-depth qualitative analysis, which is beyond the scope of this study.

As in Round 2, eight GTAs used both “Ask for Evidence” and “Ask for Work” questions; “Ask for Evidence” was used more frequently than “Ask for Work”, a reversal compared to round one. We also observed three GTAs to use “Ask about an Analogous Situation” during Round 2 (no GTAs used this question-type in Round 1).
Overall, we found that after receiving feedback, more GTAs used Stretch-It questions. Additionally, both the quantity of questions and the variety of question types increased. We infer that rehearsal in the simulator with the opportunity to observe peers and receive personalized feedback supported GTAs to implement Stretch-It questions.

**GTAs Continue To Use Stretch-It in Subsequent Simulator Sessions**

We report GTAs’ use of Stretch-It questions during all four simulator sessions in the high-intensity training semester. This allows us to compare GTAs’ performance before and after the intervention on Stretch-It. Counts (turn-of-speech) for each sub-category of Stretch-It are displayed as stacked bar charts in Figure 8. Black boxes represent Physics I GTAs and gray boxes represent Physics II GTAs.

Out of all four sessions, the total turns of speech a Stretch-It question was posed by all GTAs was found to be the largest (55 turns-of-speech) in the session focused on asking Stretch-It questions (session two), and the smallest (13 turns-of-speech) in the session focused on cold call and error framing (session one) This finding is expected since session two focused specifically on questioning skills, and session one occurred before the GTAs practiced questioning skills in the simulator.

We also observed GTAs continuing to ask Stretch-It questions during session three (group dynamics) and four (integrated skills) when the focal skill was no longer Stretch-It. The total turns of speech a Stretch-It question was posed in session three (22 turns-of-speech) and session 4 (25 turns-of-speech) were smaller compared to session two, but they were close to twice the total turns of speech in session one. The results
show that GTAs' use of Stretch-It is maintained even if they were not explicitly prompted to practice Stretch-It, suggesting an intermediate effect from the Stretch-It session.

Figure 8: Stretch-It Count across Simulator Sessions. (A) Session One; (B) Session Two; (C) Session Three; (D) Session Four. The four sub-categories of Stretch-It are as abbreviated: W – “Ask for Work”, E – “Ask for Evidence”, AS – “Ask about Analogous Situation”, AW – “Ask for Another Way to Answer”
GTAs Use All Sub-categories of Stretch-It During Sessions with Feedback About That Skill

During sessions two and four, GTAs received tailored feedback about their use of Stretch-It. We found GTAs used all four sub-categories of Stretch-It during the sessions where the facilitators gave feedback about Stretch-It (sessions two and four). This finding suggests that facilitator-led feedback and reflection was important to the GTAs’ use of Stretch-It in the simulator. For example, even though GTAs posed a Stretch-It question for a similar amount of turns-of-speech during the group dynamics session (session three) and all target skills session (session four), GTAs asked all four sub-categories of Stretch-It (at least one GTA per sub-category) during session four in contrast to session three which had guided feedback pertaining to group management where GTAs predominantly used “Ask for Evidence” questions.

GTAs Might Be Using Stretch-It As a Group Management Strategy

We observed GTAs to continue to use “Ask for Evidence” during the simulator session focused on group dynamics (session three). During this session, facilitators instructed GTAs to practice interacting with student groups with various dysfunctions (e.g., one avatar-student dominated the discussion). Due to the amount of GTAs who used “Ask for Evidence” during session three, it is possible GTAs used “Ask for Evidence” as a group management strategy. For example, after hearing from Sean and Kevin that they did not work with Maria, GTA 6 attempted to encourage communication between the avatar-students by asking Maria for the reasoning behind her idea (dialogue included in Appendix B). In Study Two, one student alluded to their GTA’s use of questioning before explaining as a strategy that helped resolve their group challenge.
Investigating GTAs’ Use of Stretch-It: Explain Logic in the Classroom

Since we coded the two main Stretch-It categories during classroom observations, we present and discuss GTAs’ use of Stretch-It: Explain Logic and Stretch-It: Follow-Up separately. As described in the Reeves et al. framework, GTAs who receive the same training might have different responses to the training [27]. Here, we present the average use of Stretch-It in the classroom for each Fall 2019 observation. To contextualize the use of Stretch-It by GTAs who participated in the simulator training, we also include pre-simulator training GTAs’ use of Stretch-It in the classroom.

Average Use of Explain Logic Across GTAs

The average use of Explain Logic per GTA for individual lesson observations in Fall 2019 is shown in Figure 9. Each black dot on the graph represents the average count of two-minute intervals per GTA observed teaching that lesson in Fall 2019. Error bars represent standard errors of average count per GTA for individual lesson observations. In addition, we calculated the average per class period observed (i.e., average per GTA per lesson) in Fall 2019, represented by a short, dashed line. The orange-colored shading in Figure 9 represents the standard error of the average in Fall 2019. For comparison, we also calculated the average use of Explain Logic per class period observed and standard error from the pre-simulator training semester (Fall 2018), represented by a long-dashed line with green-colored shading in Figure 9. We do not display the average per GTA for individual lesson observations from Fall 2018 since there was no intervention in that semester.
From the graph, we infer that Fall 2019 GTAs’ average use of Explain Logic in the classroom was impacted by the training; however, the impact was not stable. The highest average use of Explain Logic in the classroom across GTAs who participated in the simulator training (Fall 2019) occurred during the observation of the Unit 4 lesson with an averaged value of 3.83 two-minute intervals across GTAs. This observation took place the week after GTAs participated in simulator training focused on using Stretch-It (refer to timeline in Figure 6). When comparing error bars for each observation, we find Fall 2019 GTAs’ use of Explain Logic while teaching the Unit 4 lesson is distinct from their use of Explain Logic when teaching Unit 7 and Unit 11 lessons; error bars for the
Unit 4 observation do not overlap with error bars for observations of Unit 7 or Unit 11. This finding suggests the simulator session focused on using questioning skills had an immediate impact on GTA questioning behavior in the classroom. However, we find GTAs to decrease their average use of Explain Logic as the Fall 2019 semester progressed. During the observation of Unit 7 and Unit 11, the averaged values were 1.83 and 1.23 two-minute intervals across GTAs, respectively. These observations took place one to two weeks after GTAs participated in simulator sessions focused on group management and using all target pedagogical skills. Therefore, we observed an immediate impact of the simulator training on GTAs’ use of questioning in the classroom, but the implementation “boost” to the skill was not stable.

Across observations, GTAs who participated in the high-intensity simulator training (Fall 2019) had a higher semester average use of Explain Logic in the classroom than the pre-simulator training cohort (Fall 2018), with averaged values of 2.27 and 1.67 two-minute intervals across GTAs in each cohort. However, both cohorts’ average use of Explain Logic across observations is not significantly unique; error bars for Fall 2018’s semester average overlap with those for Fall 2019. In addition, we tested both sets of observation data for outliers and found one high outlier in Fall 2018’s data set. We calculated the average across observations excluding the outlier observation for Fall 2018 (1.26 two-minute intervals across Fall 2018 GTAs), as shown by a similar dashed line in Figure 10. While the overlap of error bars for the high-intensity training semester average and the pre-simulator training without outlier semester average is not as large as shown in Figure 10, the averages are not distinct.
Figure 10: High-intensity Simulator Training (Fall 2019) Cohort’s Average Use of Explain Logic without Outlier Observation. Each point on the chart represents the average count of two-minute intervals Explain Logic was coded across GTAs who were observed teaching that lesson. Both lines represent the average count of two-minute intervals Explain Logic was coded across observations for pre-simulator training (Fall 2018) and high-intensity simulator training (Fall 2019) semesters. Fall 2018’s average does not include the outlier observation. Shading represents the standard error of the mean (as error bars) for each semester average.

Fall 2019 Observation of Unit 4 Unique To Fall 2018 Semester Average

High-intensity simulator training GTAs’ average use of Explain Logic while teaching Unit 4 is distinct from pre-simulator training GTAs’ average use of Explain Logic across observations for Fall 2018. This is not the same conclusion when comparing GTAs’ use of Explain Logic for the other two observations in the high-intensity training semester. This finding indicates that while high-intensity training GTAs’ use of Explain Logic averaged across observations is not different than pre-simulator GTAs’ use across averaged observations, GTAs’ average use of Explain Logic after
simulator training is different than pre-simulator GTAs' use of Explain Logic. Therefore, the simulator training likely impacted GTAs' use of Explain Logic in the classroom.

**GTA’s Individual Use of Explain Logic**

When investigating average difference in implementation of Explain Logic for both cohorts, we noticed large error bars for both semester averages and Fall 2019 observation averages. Reeves et al. suggests GTAs who receive the same GTA training might implement different pieces of the training in their classroom [19]. Here, we chose to investigate individual results.

To represent the distribution of Stretch-It use by both cohorts of GTAs, we created heat maps, Table 4 to display two-minute interval counts for individual observations conducted in pre-simulator training (Fall 2018) and high-intensity training (Fall 2019) semesters. We operationalized GTAs’ use of Explain Logic during classroom observations (level of implementation) by categorizing ranges of coded two-minute intervals: none (0), low (1 – 3), medium (4 – 6), and high (>7).

While similar percentages of GTAs from both cohorts used Explain Logic questions in the classroom (75% and 77%, respectively), more GTAs asked a high level of Explain Logic questions in the high-intensity training semester. Five GTAs in Fall 2019 used a high-level of Explain Logic while only one GTA in Fall 2018 used it. Out of all three observations in Fall 2019, we found the largest percentage of GTAs with medium or high levels of Explain Logic in the first observation after the simulator training session focused on Stretch-It. Here, we notice five (out of seven) GTAs in the Fall 2019 cohort with medium or high levels of implementation during the first observation taught
sections of the Physics II mini-studio. We discuss implications of this finding in the Discussion section.

**Table 4: GTAs’ Use of Stretch-It: Explain Logic in the Real Classroom.**

<table>
<thead>
<tr>
<th>GTA</th>
<th>Observation by Lesson</th>
<th>Average Level of Implementation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unit 4</td>
<td>Unit 7</td>
</tr>
<tr>
<td>I</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>II</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>III</td>
<td>-</td>
<td>3</td>
</tr>
<tr>
<td>IV</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>V</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>VI</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>VII</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>VIII</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

In the table above, shading indicates a different level of implementation of Explain Logic in the classroom (range of two-minute intervals, shade): none (0, white), low (1 – 3, light gray), medium (4 – 6, dark gray), and high (>6, black). A dash indicates the GTA did not have an observation that week. *Partial Data Collected. †Outlier observation.
Additionally, we calculated individual semester average use of Explain Logic for GTAs in each cohort. We find the spread of GTAs’ semester averages of Explain Logic use in the classroom to be qualitatively different across cohorts. In Fall 2018, one (13%) GTA had a high average use of Explain Logic, while the remaining GTAs had low (N=5, 63%) or no (N=2, 25%) average use. In Fall 2019, while no GTAs had a high average use, four GTAs had medium average use (28%), five had low average use (36%), and three had no use (21%) using Explain Logic. This finding implies training focused on questioning skills influenced more GTAs to use higher levels of Explain Logic in the classroom.

*No Measured Impact of Training on GTAs’ Use of Follow-Up*

Unlike Explain Logic, a higher percentage of GTAs from the pre-simulator training cohort used Follow-Up in the classroom than the high-intensity training cohort (62.5% compared to 46.1%). In addition, we notice when a GTA from either cohort used Follow-Up in the classroom, they used it at a low-level and often used Explain Logic in the same observation (Individual GTA use of Follow-Up presented in Table 11 in Appendix D). Furthermore, when conducting an outlier test, we found all observations from both sets of classroom data with at least one two-minute interval coded for Follow-Up were considered high outliers. These findings imply that GTAs, regardless of training, infrequently used Follow-Up in their classroom.
Discussion and Implications

GTAs Benefit from Facilitator-led Feedback and Reflection

We found GTAs used all categories of Stretch-It when facilitators guided GTAs’ reflection and provided specific feedback about their use of Stretch-It in the simulator. While this provides strong evidence to suggest feedback was important to the GTAs’ transfer of the skill to the classroom, we did not disentangle feedback from practice when investigating the impact of the training on GTAs’ behaviors in the classroom. However, we observed GTAs to decrease their use of Explain Logic in the classroom over time, similar to the finding by Becker et al. (2017) [19]. The combination of these two findings imply that a deliberate practice model might be crucial to GTAs’ implementation of evidence-based practices during training sessions and in the classroom. We suggest to GTA training developers and facilitators to provide opportunities throughout the semester to practice using pedagogical skills outside of the classrooms while continuously asking GTAs about their use of target skills in the classroom during training sessions.

Follow-Up Questions Might Be Difficult To Ask

We did not find a direct impact of the simulator training on GTAs’ use of Follow-Up in the classroom. Overall, Explain Logic was used more frequently than Follow-Up during both simulator sessions and classroom observations. One possible reason might be that GTAs found Explain Logic questions less difficult to ask than Follow-Up questions, as suggested by GTA 10 during simulator session two.
GTA 10: “Um so some of them [questions] seem like things I do more naturally 'cause it will actually be about something they did or something, but things like asking it, to integrate it, asking them to integrate it to a related skill or like the new setting like those require more thought… Some of the earlier things on this list are easier to come up with but like I want to try to fit in something like with like integrated skill or like applied setting."

While we provided GTAs with examples of how to ask Follow-Up questions in the simulator, it is possible GTAs needed more examples and time to become proficient in using these types of questions in the simulator and their real classroom. Since asking about analogous situations and for other ways to explain supports higher order thinking, we suggest providing GTAs with more examples of Follow-Up in addition to asking GTAs to think of analogies to common themes in their physics course.

Some GTAs Reported Student Frustration with Stretch-It

As informed by Reeves et al., we expected higher levels of implementation of Stretch-It by GTAs in the classroom as training directly impacts GTA cognition and teaching practices [27]. During reflection and feedback sessions, we asked GTAs to think about their performance of the skill in the simulator and how they might apply it (or are applying it) to their interactions with real students. Two GTAs voiced their concern of using Stretch-It in the classroom.
The Case of GTA 13

GTA 13 was responsive to the feedback provided by facilitators and shifted their use of the skill during session two. After their final simulator round for simulator session two, we asked GTAs in that session if they considered using Stretch-It in their real classroom.

Facilitator Geraets: “Do you think you could use it [Stretch-It questions in the classroom]?“

GTA 13: “Yeah actually I noticed something yesterday too. We had this worksheet that compared the fans, the airfield to the electric field and it really helped them. So, this is kind of similar that you can apply to more familiar situations and compare it with, it really helps. It’s something I will try more.”

Thus, GTA 13 found Stretch-It questions, in particular Follow-Up, to be useful for student learning since they observed students were supported by similar question-types in the tutorial worksheet. In addition, we observed GTA 13 to use both types of Stretch-It during observation one.

However, we did not observe GTA 13 to use Stretch-It in the classroom after observation one. During simulator session four, we asked GTAs about their goal for the session, and GTA 13 stated, “Yeah. I will probably focus on the Stretch-It things because I don’t know, sometimes it annoys them, but I know it’s very, very important.”

Thus, GTA 13 still believed Stretch-It was useful for student learning, but their students’ response as annoyance led them to stop using Stretch-It in the classroom.
The Case of GTA 1

GTA 1 shared a similar experience as GTA 13 with student frustration associated with asking Stretch-It questions. After round two during simulator session two, we asked the GTAs if they would consider using Stretch-It in the classroom.

Facilitator Saitta: "How do you think this would look like to do in your classes that you TA for?"

GTA 1: "(laughs) I tried it now [Stretch-It] and they get frustrated with me."

Facilitator Saitta: "What do you mean? How so?"

GTA 1: "I'll try to get them to see what they are doing wrong but then they are just like 'I don't know what I am doing so I can't tell you what I am doing wrong'.

Facilitator Saitta: "Okay. So, I would continue to keep asking them because that is a good thing for them to just continue to think…"

While GTA 13 stopped using Stretch-It in the classroom, GTA 1’s use of Explain Logic did not fit any trend. In fact, we observed GTA 1 to use Explain Logic with high level of implementation in the classroom during observation three, despite their mention of continued student frustration mentioned before their first round in simulator session four, "I usually do a pretty good job with stretching it. I just try, even if they get it right, I ask them 'why is that? Tell me.' My students get mad at that because [they are like] ‘why are you doing this if I am right?’ (small laugh)."
It appears GTA 1 was already using Stretch-It questions in their classroom before simulator session two, as mentioned by GTA 1 and from our simulator session findings (GTA 1 asked for evidence for one turn of speech during simulator session one). Perhaps GTA 1’s familiarity with Stretch-It before focused rehearsal on the same skill contributed to their resilient use of Stretch-It despite student resistance. It is also possible that GTA 1 valued asking questions in the classroom.

Based on the cases of GTA 13 and GTA 1 we infer that although GTAs might experience similar situations while implementing Stretch-It in the classroom, each GTA might have a different reaction to student resistance. The curricular shift to support inquiry and sense-making might be unexpected for students who expected their physics course would be non-interactive [124]. The varying cases of GTA 13 and GTA 1 highlight a flaw in current GTA professional development with respect to how to react to student resistance. As suggested by Reeves et al., GTA characteristics, including teaching experience, influence GTA beliefs about teaching and their teaching practices [27]. Thus, if a GTA has a negative experience with using a pedagogical skill in the classroom, they might associate their student’s dissatisfaction with that skill, possibly causing GTAs to not use the skill again. This finding calls for further investigation into how and why GTAs may or may not implement pieces of professional development into their classroom.

Furthermore, we did not discuss ways GTAs can manage student resistance or frustration in the classroom, even though we were introducing GTAs to pedagogical skills that might violate how students expect their physics GTA to interact with them. We strongly suggest to GTA training developers and facilitators to incorporate discussions
focused on strategies to manage student resistance with respect to evidence-based teaching practices and active learning instruction with GTAs.

**GTAs Create a Culture within Their Prep Meeting Cohort**

During the Fall 2019 observation of the Unit 4 lesson, we noticed most GTAs who used Explain Logic a medium or high amount taught Physics II mini-studio. We suggest two possible reasonings for this finding. Firstly, it is possible the Unit 4 lesson for Physics II mini-studio supported GTAs' use of Explain Logic, implying that the curriculum might have influenced their use of the skill. Another possible reason is that the GTA cohorts for Physics I and II created their own cohort culture which varied in receptiveness to the training. As previously mentioned, two researchers (T.W. and C.M.D.) lead separate weekly preparation meetings for Physics I and Physics II GTAs in the Fall 2019 cohort. Both researchers noted differences in how the two cohorts interacted. Physics II GTAs seemed interested in helping each other with understanding the worksheet/lab, while Physics I GTAs did not demonstrate the same support system. However, we find it difficult to disentangle the effects of the curriculum and teaching culture. Future work will focus on investigating these variables and the potential impact each one might have on GTA teaching practices.

**Limitations**

We chose to live code instead of use video-analysis to document GTA behaviors in real-time; this allowed us to collect data for more GTAs. Similar to the equipment used by tour guides, we used a transmitter and receiver set to listen to GTA-student
conversations during classroom observations. GTAs wore a lapel microphone hooked up to an audio transmitter while researchers listened to GTA-student conversations through headphones attached to a receiver. An audio-recorder was fed into the loop to record. However, we encountered issues that might have led to an observer missing a code on a two-minute interval. Examples include audio interference, wire connections coming loose in the middle of an observation, and difficulty hearing over noisy student groups. Due to the complex nature of live coding, it is possible Stretch-It is underrepresented in our findings. We also discuss this limitation in previous studies [50,116].

In our study, we explored the impact of simulator training with physics GTAs from one institution whereas Reeves et al. states it is vital to investigate the influence of the contextual variables, including investigating similar training across institutions (like [65]). Future research would include exploring the use of a simulator for training physics GTAs who teach in student-centered recitation and lab sections across institutions. Also, we did not disaggregate our findings by many GTA characteristics, which could affect both GTA receptiveness to the skills and student response to GTA use of the skills.

In addition, the mini-studio sections are not stand-alone courses and were paired with lecture sections; mini-studio grades counted an average of 20% of the lecture final grade. It is possible GTAs were influenced by the lecture instructor to implement different teaching practices or activities during either the recitation or lab portion (i.e., during one observation for GTA 1, GTA 1 went over a practice test with students after they asked what students wanted to work on during recitation, the practice test was
taken from the lecture course website), as was demonstrated in a multi-tiered professional development project for calculus instructors and GTAs [125].

As part of our discussion, we briefly mention GTAs who taught different courses might have created their own culture of teaching based on trainer observation during weekly preparation meetings. We also recognize each lesson might provide GTAs with different opportunities to use Stretch-It in their mini-studio section. However, we find it difficult to disentangle the culture and the lesson from our findings. Future work will aim to incorporate these two pieces into our investigation.

**Study Conclusion**

In summary, we observed GTAs to increase or shift their use of Stretch-It across rounds during the session focused on using Stretch-It. In addition, we observed all four sub-categories of Stretch-It to be used by GTAs during simulator sessions with focused feedback and reflection about using Stretch-It during their practice rounds in the simulator. We also observed most GTAs to use “Ask for Evidence” during the group dynamics simulator session suggesting GTAs might have used that sub-category of Stretch-It as a strategy for group management.

In addition, we found simulator training focused on the use of questioning to impact GTAs’ use of questioning in the classroom. Specifically, we observed more GTAs to use a higher rate of Explain Logic questions during the classroom observation immediately following practice of Stretch-It in the simulator. We also observed more GTAs in the high-intensity training semester to use a higher rate of Explain Logic questions than GTAs in the pre-simulator semester which suggests more GTAs are
intentionally using Explain Logic in the classroom because of the training. However, the simulator training did not impact GTAs’ use of Follow-Up in the classroom; GTAs might randomly use Follow-Up because of the curriculum or other factors such as GTA characteristics or cognition.

**Future Work**

It is possible for GTAs to elicit student reasoning by asking “Why?” but not value alternate student ideas by guiding students to a physically correct answer [10]. Future work would include investigating the intention of different types of questions asked by GTAs in the classroom.

With regards to the varying responses of student frustration with using Stretch-It in the classroom, future work will include discussions about student resistance to using active learning strategies in the classroom and how to mitigate negative responses to evidence-based practices. In addition, while we did not collect demographic information from GTAs, future work will incorporate pieces of identity such as gender and ethnicity.
CHAPTER SEVEN: STUDY TWO – STUDENT PERCEPTIONS OF GTA STRATEGIES TO MANAGE GROUP CHALLENGES

Introduction and Background

Tutorial and student-centered lab environments provide students with opportunities to collaborate with peers while learning physics. During tutorial, students are expected to have discussions as they try to make sense of physics and everyday thinking [7]. Conlin and Scherr argued a safe space to “sensemake” must be established within the student group to achieve the optimal benefits of tutorials on student learning [59]. However, their study mainly focused on student behaviors and did not investigate how GTAs might influence group dynamics.

In lab, students are expected to work together to perform experiments while focusing on scientific research processes used by physicists [8]. Similarly in lab spaces, it is crucial for students to create a safe space to productively reason through the inquiry process like “problematizing” [60]. Furthermore, social dynamics of the group might influence student engagement during inquiry-style labs.

From these descriptions, tutorial and lab settings heavily rely on social interactions between students. In these situations, students might feel vulnerable when sharing their ideas with their peers or their instructor. Downing et al. found community college students enrolled in an active learning-style biology course reported mixed feelings of anxiety associated with group work [47]. One of the reasons some students felt group work increased their feelings of anxiety is due to fear of negative evaluation [47].
Given the large amount of tutorial and lab sections compared to lecture sections, graduate teaching assistants (GTA) often teach these reformed courses and might have an impact on student group dynamics. While group management might be discussed in STEM GTA training, GTA-student interactions observed in such spaces have been limited to classroom discourse. For example, Goertzen, Scherr, and Elby explored how GTAs “look for indicators” of student learning and the reasons why GTAs use such behaviors [109]. West et al., described GTA-student interactions using a protocol that documents codes like open dialogue or closed dialogue [12]. However, such protocol might not be applicable for investigating group management strategies used by GTAs as the group issue must be resolved in the perspective of the student group. In our previous study, we similarly found different GTAs to use different teaching styles in mini-studio (combined tutorial and lab) sections, including a non-interactive instructional style [50].

Furthermore, strategies physics GTAs use to help students feel comfortable sharing their ideas in front of their peers is limited. This study aims to contribute to the literature about how physics GTAs recognize and help students to resolve their group challenges. Implications from this study will inform GTA training efforts.

**Research Questions**

In this study, we investigate the student perspective of group work in introductory physics mini-studio spaces with regards to group dynamic challenges and the social actions their GTA took to help remedy the group’s situation. Therefore, we address the following questions:
1. What group challenges do students experience in the physics mini-studio sections?

Next, we explore student perceptions of their GTA's awareness of their group’s situation and the strategies their GTA used to help resolve the issue. Therefore, we ask:

2. How do students perceive their GTA to recognize and help resolve their group challenge?

Finally, we explored the student perception of their GTA’s role (mentor or instructor) in the mini-studio section and the relationship between that role and their perception of whether their GTA helped resolve group challenges:

3. How do students perceive their GTA’s role in the classroom, and how is it related to their GTA resolving their group challenges?

By answering these questions, we aim to contribute to the literature about how GTAs recognize and help to solve group dynamic issues. Also, we make recommendations for physics GTA training.

**Methodology**

*Using Case Studies to Explore Student Perspective of Group Management*

Following a multiple-case study methodology [105,106], we explored how students perceived their GTA to resolve their group challenges. We considered each student to be a single case since students have different perspectives of the atmosphere in their mini-studio section and GTA, even if some students had the same GTA. During analysis, we compared each student’s experience or case and discuss
common themes in our Findings and Discussion section like a case study by Sandi-Urena and Gatlin [64].

*Frameworks Utilized*

In tutorial and lab sessions, students are expected to work in groups to complete activities; however, a student might hesitate to interact with group members because of fear of negative evaluation. In this study, fear of negative evaluation motivated the investigation of group dynamics challenges experienced by students in such spaces. Guided by Reeves et al. we investigate the relationship between GTA teaching behaviors and student outcomes by exploring strategies used by GTAs to resolve student group challenges. Additionally, implications from this study will inform changes to be made to future iterations of the teaching intervention described in Chapter Four. We used facework to operationalize actions GTAs take to resolve student group challenges.

*GTA Training with a Mixed-Reality Classroom Simulator*

In Spring 2020 sixteen GTAs who taught sections of mini-studio (combined recitation and lab) as part of a two-semester course sequence for introductory algebra-based physics (College Physics I and II) participated in up to four sessions with a mixed-reality classroom simulator as part of their training. GTAs who were new to teaching mini-studio and to the simulator training participated in three simulator sessions focused on specific pedagogical skills, and a fourth session focused on integrating all the target skills into their rehearsal. Returning GTAs participated in up to
three sessions, and each session was structured like the fourth session for new GTAs. Nine out of sixteen GTAs participated in the training for both professional development and research purposes. A timeline for Fall 2019 and Spring 2020 is presented in Chapter Four, Figure 5.

Participants

Fifteen undergraduate students enrolled in either a Physics I or II mini-studio section participated in the semi-structured interviews about their interactions with their GTA and their experience with group work in their mini-studio section. The GTAs varied in their amount of experience teaching the mini-studio course and participation in simulator training. Table 5 indicates the course in which each interview participant was enrolled, a code for their GTA, and their GTAs’ level of teaching experience.

Table 5: Undergraduate Physics Student Participants

<table>
<thead>
<tr>
<th>Student</th>
<th>Enrolled in Physics I or II</th>
<th>GTA’s Level of Experience Teaching Mini-Studio Course</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>I</td>
<td>New/ Taught other</td>
</tr>
<tr>
<td>B1</td>
<td>I</td>
<td>New</td>
</tr>
<tr>
<td>B2</td>
<td>I</td>
<td>New</td>
</tr>
<tr>
<td>C1</td>
<td>I</td>
<td>Returning</td>
</tr>
<tr>
<td>C2</td>
<td>I</td>
<td>Returning</td>
</tr>
<tr>
<td>D</td>
<td>II</td>
<td>New/ Taught other</td>
</tr>
<tr>
<td>E1</td>
<td>II</td>
<td>New</td>
</tr>
<tr>
<td>E2</td>
<td>II</td>
<td>New</td>
</tr>
<tr>
<td>F1</td>
<td>II</td>
<td>Returning</td>
</tr>
<tr>
<td>F2</td>
<td>II</td>
<td>Returning</td>
</tr>
<tr>
<td>G1</td>
<td>II</td>
<td>Returning</td>
</tr>
<tr>
<td>G2</td>
<td>II</td>
<td>Returning</td>
</tr>
<tr>
<td>H1</td>
<td>II</td>
<td>Returning</td>
</tr>
<tr>
<td>H2</td>
<td>II</td>
<td>Returning</td>
</tr>
<tr>
<td>I</td>
<td>II</td>
<td>Returning</td>
</tr>
</tbody>
</table>
Developing Group Challenge Scenarios

In Fall 2019, we interviewed 14 physics undergraduate students enrolled in the target courses about their experiences with group work in their Physics I or II mini-studio section. After conducting the interviews, we noticed students described social challenges within their group, like not everyone contributing to the group activity.

However, we also found some students expressed challenges outside of the group’s function, like lecture and lab not being in-sync. To elicit student ideas about their group dynamic challenges, we developed three scenario prompts including topics students brought up from the Fall 2019 interviews and group dynamic issues assigned to avatar-student groups in the simulator training. Below are the three group challenge scenario prompts:

Scenario One: Miguel is working in a group with three other students on a lab activity. Miguel has assumed his role as group leader since his ideas have always been correct from the start of the semester. Everyone in the group seems to go along with his idea and each contributes to the discussion about the experiment except for one student, Sophia. Sophia remains quiet during the discussion. When it is time to conduct the experiment, she writes down everything the group says in their lab report or lab notebook. She also does not touch any of the lab equipment since Miguel is always the first one to start setting up the experiment. After the lab is finished, Sophia looks frustrated.

Scenario Two: During recitation, Ray is working with two other students on a worksheet about energy. After working through the first section, Ray finds themself confused about something on the second page. Ray has also noticed that it has been about twenty minutes and no one in their group has said anything to each other. Ray looks over and sees one student is far ahead in the worksheet and the other student is on their phone.

Scenario Three: Liam, Emily, and Charlie all have different ideas about the relationship between force and motion. Liam thinks his idea is correct and will not listen to what Emily and Charlie have to say. Emily and Charlie split off into their own group of two while Liam works alone. However, Emily and Charlie eventually become confused. Seeing that Liam has worked through the entire activity, they
decide to go along with Liam’s ideas because they do not want to get a bad grade.

**Undergraduate Student Interviews**

Towards the end of the Spring 2020 semester, three researchers (T.W., A.G.G. and C.M.D.) interviewed fifteen undergraduate students enrolled in either a Physics I or II mini-studio section taught by a GTA who participated in the simulator training. Each student individually participated in a semi-structured interview about their interactions with their GTA and their experience working in a group with their peers during their mini-studio section. Due to the COVID-19 pandemic, each interview was remotely conducted using video-conferencing software (i.e., Zoom) to ensure the safety of participants and interviewers [126].

During each interview, students were presented with the three different group scenarios. We asked students if the scenario was familiar to them in their experience with group work in their mini-studio section. Additionally, interviewers asked if their GTA recognized the situation and what strategies their GTA used to help resolve the group challenge. The student interview protocol is available in Appendix E. The interviews were audio-recorded and transcribed.

**Data Analysis**

After collecting the student interview data, we used a mixture of a priori coding and thematic analysis to analyze the student interviews [117]. To investigate group challenges not descriptive of the scenario prompts reported by students, one researcher (C.M.D.) reviewed the student interviews for examples of other group challenges from
each interview and compared the responses to create common themes. Afterwards, researchers (C.M.D. and J.C.C.) discussed and collapsed the themes into three distinct codes described in the Findings Section of this study. Then, a code book was created including the three scenario prompts, alternative group challenges (described below in the Findings Section), and GTA acknowledgement of group challenge.

**Investigating Inter-Rater Reliability**

Two researchers (C.M.D. and J.C.C.) investigated IRR for each code from the code book in two stages. During stage one, both researchers used the code book to code two (out of 15) interviews. After individually coding the two interviews, both researchers compared coding and discussed until agreement. Minor changes to the code book were made before the next stage of coding. In stage two, the same researchers individually applied the code book to the rest of the interviews (thirteen). Gwet’s AC1 was used to calculate a value that can be interpreted for investigating IRR [119]. The grain size for each code category was a student interview. Gwet’s AC1 values ranged from 0.76 to 1 across codes before discussion. After discussion, Gwet’s AC1 values for all codes were 1. A value between 0.61 and 0.81 can be interpreted as moderate agreement while a value greater than 0.81 can be interpreted as near perfect agreement [120].

**Findings and Discussion**

Here, we present student experiences with group challenges in their mini-studio and their perception of their GTA’s awareness of their group’s issues 1) as described by
the group prompts, and 2) not described by the group prompts (also referred to as alternative group challenges in this study). Then, we compare the student perspective of their GTA’s role (mentor or instructor) in the classroom with their experience with group challenges and their GTA’s awareness of the challenges.

Students Experienced Group Challenges Described by the Scenario Prompts

We find more than half (nine out of fifteen) of the students reported experiencing a group challenge similar to at least one scenario prompt. Additionally, two students expressed familiarity with more than one scenario prompt. Table 6 shows which scenario prompts students reported to be descriptive of their group challenges and the shading color refers to the student perspective of their GTA recognizing their group’s situation. Here, we present examples of student responses that describe the portions of the scenario prompts students felt were familiar.

Scenario One was the most frequently reported scenario by students (six students) to be descriptive of their group’s challenge. In short, Scenario One described a group situation where “Miguel” dominates the group’s decisions while “Sophia” remains quiet and seems frustrated. Some students felt the character of Sophia was similar to a disengaged or non-contributing member in their group. Student B2 explained the similarity, “Sophia was similar to a student that we, a lab mate that we had, who kind of just remained quiet in the lab, and didn’t really do too much unless she was asked to.” Another student, Student B1, felt the role of “Miguel” in scenario one was more descriptive of their situation. They explained, “When we would get there [lab], after we did the quiz and finish the quiz, we would, she would already automatically, like,
grab the computer start setting up and like going to the report or opening up the graphic stuff in the sensor.” Student B1 also described how their group member who assumed the role of leader created a tense environment. They stated, “It was always a whole frustration with her, you know, because she would always take it to the next level. And we’re like, ‘Dude, it’s gonna get done’. And like, everyone’s together and like, we would look around, like all other groups will be like laughing with each other and having fun and building things.”

Table 6: Students Report Experiences Similar to the Prompted Group Scenarios

<table>
<thead>
<tr>
<th>Student</th>
<th>Experienced Scenario During Mini-Studio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Scenario One</td>
</tr>
<tr>
<td>A</td>
<td>X</td>
</tr>
<tr>
<td>B1</td>
<td>X</td>
</tr>
<tr>
<td>B2</td>
<td>X</td>
</tr>
<tr>
<td>C1</td>
<td></td>
</tr>
<tr>
<td>C2</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td></td>
</tr>
<tr>
<td>E1</td>
<td>X</td>
</tr>
<tr>
<td>E2</td>
<td></td>
</tr>
<tr>
<td>F1</td>
<td></td>
</tr>
<tr>
<td>F2</td>
<td>X</td>
</tr>
<tr>
<td>G1</td>
<td></td>
</tr>
<tr>
<td>G2</td>
<td></td>
</tr>
<tr>
<td>H1</td>
<td></td>
</tr>
<tr>
<td>H2</td>
<td>X</td>
</tr>
<tr>
<td>I</td>
<td></td>
</tr>
</tbody>
</table>

In the table above, an “X” indicates the student reported the scenario prompt to be descriptive of their group dynamic challenge. Cell shading refers to the student perception of their GTA recognizing their group challenge. (Gray – GTA did not recognize the group challenge; Green – GTA recognized and helped to resolve the group challenge)
Scenario Two, where all students in the group were either disengaged from or working on different parts of the tutorial, was relatable for some students (four students). Student E2 described their group’s situation, “I would work on the worksheet before the class started, like I would work on it at home. And in the times that my classmates, or my group mates did attempt to work on the recitation, they would see that I was already done with them. They got stuck on a question. Basically, I would just hand them over my recitation and be like here, just, just look at the answer here.” In Student E2’s case, their group did not always work together (or on) the tutorial worksheet, and often times Student E2 would be ahead of the group. Student H1 also felt scenario two was familiar to their situation, but for a different reason. They explained, “I’m not [a] physics genie by any means, but it wasn’t so much, you know, one person was more ahead than the other. In this case, it was the person that was constantly on their phone, you know, just coasting by off of us.” Like Student B2, they felt their situation was familiar to scenario two because some group members would not contribute to the group work.

A couple of students felt scenario three was familiar to their situation in mini-studio. Student E2 expressed, “Scenario Three, on the other hand, is also pretty similar, I would say, because not like the whole different ideas thing, but like the people who don’t engage, or see that [what] we’re doing is right.” Again, like Student B2 and H1, Student E2 mentioned a lack of disengagement or contribution from other students to be the underlying reason why one of these scenario prompts seems familiar.

Our findings imply students enrolled in a physics mini-studio section might experience group dynamic challenges as described by the group scenario prompts and
in particular the lack of contribution by everyone in the group. It is possible that if a student experiences these group challenges the impact of implementing student-centered instruction on student learning might be diluted. For example, students that have difficulties with group work might think group-related activities are not supportive of their learning resulting in reduced student buy-in for student-centered instruction [127].

In addition, we noticed three students (Student B2, Student C2, and Student E2) talked about assignment grades in their discussions about the scenario prompts and non-contributing group members. When discussing scenario three, Student E2 mentioned, “They’ll bounce off our grade. They get a good grade, even though they didn’t do work.” From their response, Student E2 brings up their concern about uneven workload as Student E2 perceived themself to have contributed more to the assignment than the other group members despite everyone receiving the same grade. Similarly, Student B2 expressed feeling a dislike for their group’s situation because it was unfair that the Sophia-like student from scenario one did not contribute as much as the other group members. They stated, “I didn't really like that, you know, I feel like because we're all getting the same grade, we should be equal effort, and more participation.” Since getting “good” grades are a motivational factor for students to complete their course work [128], it is possible when they experience a negative group work situation, that it will increase a student’s feeling of anxiousness in the classroom. This is evident when Student C2 described their feelings about the similarity between their group challenge and Scenario Two, “I was definitely anxious just because it’s obviously my grade. And I want to do well, and I want like, [pause] I like, it was in the end my grade.” Student C2 continued discussing the reason why they felt anxious about doing more of the work:
“So, I was anxious and trying to get it, right, because obviously, you needed to get it right to do it. But also, kind of anxious because it was their grade as well. And even though they weren’t contributing, I felt like [pause] that I was like, I could have been doing it all wrong, or like, the two of us could have been doing it all wrong, and they wouldn’t have mentioned anything, and we would have just been practicing like wrong, wrong way to understand physics.” Thus, it is possible group challenges caused by poor group dynamics like having a non-contributing group member can exacerbate negative student emotions like fear of negative evaluation.

GTAs Might Not Recognize Group Dynamic Challenges

Of the eight students who expressed a similarity between their group’s issue and a group scenario prompt, six students reported their GTA to not recognize their group’s challenge. Two students did not discuss if their GTA did or did not recognize their situation described by the scenarios during the interview. Table 6 shows which scenario prompts students reported to be descriptive of their group challenges and the shading color refers to the student perspective of their GTA recognizing their group’s situation. This finding implies that most of the time when students who experienced a group challenge similar to the scenario prompts perceived their GTA to not notice their situation.

In the simulator training, we made the dysfunction in the avatar-student groups explicit. For example, Kevin would say something like, “Sean and I agreed with each other, but Maria did not” indicating there is a disagreement within the group. However, according to Student F2, their group’s challenge similar to Scenario One might not have
been so apparent. When asked if their GTA could do anything to help resolve the issue, after stating their GTA did not recognize the situation, Student F2 responded, “No, I don’t think, I mean to her it probably looked like we were all do, working together. Because he would, he would watch us do the experiment. It’s not like he was just texting on his phone not paying attention like he would engage he just wouldn’t talk or contribute, but he was engaged.” This implies that it is possible GTAs might miss group challenges when observing student groups because of the subtly of the issue in the classroom.

On the other hand, Student H2 reported their GTA to recognize and help resolve their group’s issue similar to scenario one. First, Student H2 discussed how their GTA noticed that one group member tended to leave early, as they said, “I think it was our first or second lab. And she’s like, ‘where’d she go?’ And we’re like, ‘No, we were done. We told her like, she could go, but we just wanted to make sure like, everything on the lab look good.’ So, she’s like, ‘okay, ‘cause, you know, this has to happen’. And we’re like, ‘yeah’, and she’s like, ‘you know, I noticed she likes to run out’ and I was like ‘Yeah’.” Then, the student talked about the GTA offering different solutions to the group. Student H2 said, “So [GTA] did end up telling us because one of our lab mates after lab, that person left and we’re still there. And, um, she was like, you know, I wanted, one of the group members wanted to switch it. I was like, ‘No, I like our little girl group. You know, we can talk to each other because we’re obviously adults’. So [GTA] said, you know, she said, ‘the only way someone knows what they’re doing is wrong, is until you tell them.’ But um, we didn’t want to feel like we were on ganging up on her.” Finally, Student H2 explained that the GTA discussed the role of groups with the whole class,
which helped the group resolve their issue, “So, um, [GTA] knew that there were some other groups like that other group, but that one really quiet girl. And there was some other groups who had their own problem. So, I think it was by the second lab, I my second lab is resolved. So [GTA] ended up the next week in [recitation] said, you know, guys, you know, I've noticed that there are some people, you know, who, X, Y, and Z, and she's like, you know, our group, this is what it's for, you know, you have to communicate and stuff. So, she kind of threw it out there. So, we didn't have to really pull her aside because I didn't really want to. [GTA] told us like, you know, whenever you guys got to class, you just ask me to step out and talk to her. But I feel like that's really embarrassing. If everyone sees a group go outside, they know, obviously, something's wrong. So, I'm kind of glad that [GTA] kind of pitched it because then after that, one of the group members kind of initiated the conversation there. And, um, the girl ended up getting it so after that, there wasn't really a problem.”

In Student H2’s case, their GTA attempted different strategies before finding one that helped resolve the group’s issue. It is possible that the other students’ GTAs attempted to resolve group management challenges but did not continue trying to help after their first strategy failed. For example, when asked if their GTA noticed their group challenges, Student B2 responded, “No, I don't think so. Because there's never a change. Only thing is like midway into the lab, about like, mid-February, she did offer for us to change groups. So, I don't know if maybe that was what was behind it.” However, Student B2’s GTA did not change the groups even though Student B2 thought that might help resolve their group challenge as they expressed, “Yeah, I wish she would
have just automatically switched us up, because that could have been pretty helpful and nice.”

**Students Experienced Alternative Group Challenges**

To investigate other group dynamic challenges experienced by students, we asked if they experienced other challenges in their group not described by the scenario prompts. We find eight out of fifteen students reported experiencing at least one of three alternative group challenges not described by the group scenario prompts.

The most commonly student-reported alternative group challenge was “all group members feel stuck”. Some (five) students expressed their group to become “stuck” or “lost” when working on an assignment to the point where students could not proceed. For example, Student G1 described their group’s challenge during lab, “One issue I might think of with my group is um when we were when we were working through the labs and doing the experiments. And I mean, this, this is kind of like a, like a standard thing. But we would really become stuck with stuck on one of the problems, we’re really not sure how to proceed.”

Another common alternative group challenge described by students was “disagreed on how to solve problem”. A few students mentioned their group would have disagreements about how to solve or proceed with a problem. Student B2 explained their situation during the quiz, “But in actual lab, like face-to-face lab, it was more of like during the quiz, which also I feel like we should have taken individually, because people
had different opinions about what to do, and how to approach the question. So um, like, we did get into a little bit of a squabble with that.”

**Student Frustration with Tutorials Might Influence Group Members to Not Contribute**

Expressed by one student was the alternate group challenge “group members do not equally contribute or share workload”. Student E2 exclaimed, “It’s similar to Scenario Two, but it’s not exactly like Scenario Two, where it’s like, I know what scenarios two’s asking the lack of teamwork and engagement within the group. So, I would say that.”

This was an underlying theme discussed by other students who felt the scenario prompts were descriptive of their group challenges. While Student E2 felt Scenarios Two and Three were descriptive of their group challenges, they also described this alternate group challenge to reflect the overall classroom environment. Throughout the interview, Student E2 expressed dissatisfaction with tutorial and a lack of engagement from their GTA resulting in their own disengagement from the course. When asked why they felt this way, Student E2 said, “So, I was like, at first, I was very frustrated, because I was like, well, I want a good grade in this. And this is like, going to count towards my final physics grade. So, if this is going to screw up my grade, because the TA is not engaged, then I'll be very upset. But seeing that, like, I was able to do well in the lab reports for the most part, and that he didn't grade strictly, that I was able to, like the recitation worksheets weren't graded. So, I just stopped doing them because I got disengaged. And like, I was doing well on the quizzes, my group and I, so I have I have an A in the lab. So, I was like, Okay, I really don't have to worry at this point.”
An alternative way of thinking about Student E2’s experience is that their GTA stopped interacting with Student E2 because of unsuccessful attempts to engage with or motivate Student E2 to work on the tutorials. From Student E2’s response, they did not find the tutorials to benefit their learning, and therefore they chose not to participate. This is not an uncommon finding when students do not have “buy-in” for reformed curricula [124,127]. Similarly, in Study One, when students expressed frustration with Stretch-It, GTA 13 stopped using that strategy in their classroom despite recognizing the benefit of asking those types of questions on students learning. In both instances, GTAs disengaged because of student frustration. In Student E2’s case, their dissatisfaction with the curriculum resulted in their frustration and GTA’s disengagement. This implies GTA training should address strategies with how to manage student frustration since it is a barrier for productive group work and student learning.

Students Perceived GTAs to Help with Group Challenges Related to Physics Content

Unlike the group scenario prompts, five out of the nine students who reported an alternative group issue perceived their GTA to recognize and help with resolving their group’s challenge. One student reported their GTA to attempt to help their group solve their issue, but their strategy did not help. Another student perceived their GTA to recognize the situation but did not do anything to help resolve the issue. Two students did not discuss their GTA’s awareness of one or more of the alternative group challenges they reported. Table 7 shows which students expressed the emergent group
challenges, indicated with an “X”, and the student’s perception of their GTA’s awareness of the group’s issue (indicated by cell shading).

Table 7: Students Report Group Challenges Not Described by the Prompted Group Scenarios

<table>
<thead>
<tr>
<th>Student</th>
<th>Other Challenges Experienced During Mini-Studio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Group members not equally contributing or sharing workload</td>
</tr>
<tr>
<td>A</td>
<td>X</td>
</tr>
<tr>
<td>B1</td>
<td></td>
</tr>
<tr>
<td>B2</td>
<td></td>
</tr>
<tr>
<td>C1</td>
<td></td>
</tr>
<tr>
<td>C2</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td></td>
</tr>
<tr>
<td>E1</td>
<td></td>
</tr>
<tr>
<td>E2</td>
<td></td>
</tr>
<tr>
<td>F1</td>
<td></td>
</tr>
<tr>
<td>F2</td>
<td></td>
</tr>
<tr>
<td>G1</td>
<td></td>
</tr>
<tr>
<td>G2</td>
<td></td>
</tr>
<tr>
<td>H1</td>
<td></td>
</tr>
<tr>
<td>H2</td>
<td></td>
</tr>
<tr>
<td>I</td>
<td></td>
</tr>
</tbody>
</table>

In the table above, the color shading indicates if the student perceived their GTA to recognize and/or help to resolve their alternate group challenge. Green – GTA helped resolve issue; Yellow – GTA attempted to help resolve issue; Orange – GTA noticed issue but did not attempt to help resolve issue.
The four students who discussed their GTA’s strategy for helping (or attempting to help) the group with their issue related to “all group members feel stuck” described their GTA to explain to them. Student G1 explained, “…we would really not be able to proceed, like we’d just be sitting there like talking to each other, like figuring out how to proceed. And for that [GTA’s name omitted] will come over and we just asked him, like how, like what we do now. And he was like, really good pointers on how to proceed with the lab.” Student G2 described their GTA (same as for Student G1) to use the same strategy when their group got “stuck”. However, their GTA’s explanations were not always helpful and did not help resolve the issue as they expressed, “Sometimes, like the TA wouldn’t really explained clearly.”

Both students who described their GTA to helped with resolving their group challenge related to “Disagreed on how to solve the problem” reported their GTA to use different strategies. Student E1 reported their GTA’s strategy was providing explanation when students asked for help, the same strategy previously discussed. Student C1 described their GTA to listen to everyone’s idea in the group before providing hints of which student might have a correct idea. They said, “Yeah, definitely, um, he would listen to what everyone was saying to the different answers. And then he would, you know, kind of make a hint at who’s, who’s solving it the right way, who’s solving it the wrong way. And then we would go from there.”

Thus, we found most GTAs were reported by students to resolve their content-related group challenges by using explanations. It is possible GTAs might feel more comfortable with talking to students than using other teaching behaviors like those
described by Student H2 in a previous section. Wilcox, Yang, and Chini observed GTAs who taught the target courses before the teaching intervention to explain to students on average 25% of the time while they engaged in open dialogue only 4% of the time [2]. If GTAs feel more comfortable with explaining, than it is not surprising they might resolve content-related group challenges and not group challenges as a result of group dysfunction. We also acknowledge that while a GTA’s default to explaining can be helpful for some groups to move on to the next part of the activity, that strategy does not work for each student group (i.e., Student G2). This implies GTAs might want to consider using a variety of teaching strategies to support a variety of student learners which can be scaffolded during GTA training.

Moreover, we notice students enrolled in Physics II mini-studio sections reported more content-related challenges resolved or attempted to be resolved by their GTA than students enrolled in Physics I mini-studio sections. We expand on this claim in this next section.

*Student Perspective of GTA’s Role: Mentor vs. Instructor*

To understand how the environment created by a student’s GTA might be related to a student’s experience with group challenges and their GTA helping to resolve their challenges, we investigated whether students perceived their GTA as a mentor or instructor. We find four students reported their GTA to only be a mentor, four students thought their GTA was both a mentor and instructor, four students said their GTA was only an instructor, and two students reported their GTA to be something other than mentor or instructor. One student did not discuss their GTA’s role during their interview.
We sorted the students into four different groups depending on how they perceived their GTA’s role in the classroom. Table 8 shows which students were sorted into each category along with the group challenges they experienced and their GTA’s recognition of their situation.

Table 8: Relating Experience with Group Challenges And Perception of GTA

<table>
<thead>
<tr>
<th>Student</th>
<th>Scenario Prompts</th>
<th>Other Challenges</th>
<th>GTA Role</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>C1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In the table above, the color shading indicates if the student perceived their GTA to recognize and/or help to resolve their alternate group challenge. Green – GTA helped resolve issue; Yellow – GTA attempted to help resolve issue; Orange – GTA noticed issue but did not attempt to help resolve issue.
Mentors Create Comfortable Learning Environments

Here, we use student descriptions of mentor and instructor rather than pre-determined definitions to allow students’ perception of the learning environment to emerge. All student quotes describing their GTA’s role are available in Appendix F.

GTAs who were perceived by students as mentors (including mentor-instructor combination) might have created a classroom environment such that students are comfortable sharing their ideas. Student C1 supports this claim with their response to the question about whether they would consider their GTA a mentor or instructor, “Um definitely more like a mentor. I didn’t, I think we all felt very comfortable with him.” Student H2 also described their GTA to be relatable creating a comfortable classroom environment, “She’s just relatable. I don’t know. I, she just, I don’t know, she makes you feel comfortable. At least for me, she was comfortable. She’s like, ‘it’s alright. I understand.’” Being relatable can contribute to the creation of a supportive learning environment [129].

Some students also described their GTA to provide a personal or individualized learning experience when discussing how their GTA related to being a mentor. Student F2 described, “I feel like she wasn’t really just talking at us all the time. She really would take her time to speak to us individually and see how she can get the point across to each of us individually, as opposed to just like, just telling us all the facts.” Student G1 agrees with Student F2 as they mentioned in their description of their GTA, “…as far as like, on a personal level, like helping us, I would say as a mentor.” When discussing how their GTA was a mentor and not an instructor, Student H1 said, “To me, I see
mentors as you know, somebody who's looking over your shoulder and guiding you along the way and helping you out. Whereas an instructor is that person at the front of the class, you know, lecturing you to death, and it's just very unenjoyable.” From the student descriptions above, a mentor is someone who provides a personalized learning experience which is similar to “cares about your learning as an individual”, an example of facework highlighted by Gaffney and Gaffney [42].

On the contrary, when students described their GTA as an instructor, they did not express a comfortable learning environment. Student B2 exclaimed how lab should be comfortable and the GTA should be relatable, “I mean, as far as like making sure that we were very punctual with our time as far as like going from recitation, like what the worksheet to the quiz to then the lab, she definitely did pretty well. But I feel like to make her more of a mentor is to make her a little bit more relatable. You know, we already get a lot of instruction from her, regular professors. So, you know, I feel like going to lab should be more comforting and a little bit more relatable, especially because the TA is a little bit more closer in age and pretty much went through the same thing that we went through already.” It is possible students might still feel comfortable with their GTA as an instructor especially if their GTA is helpful, but they did not mention comfortability in their descriptions of instructor. For example, Student E1 mentioned their GTA was helpful, “I don't think we were not, I mean, at least, in my experience, we weren't on a personal level like that. He really helped me to understand concepts in a professional manner. Not that a mentor isn't professional, but we never like discussed our life or, like, what I'm doing in school or anything like that we solely talked only about physics in that
class. And he also did a good job at maintaining, like, good, professional professionalism, as the TA of the class, but yeah, definitely an instructor.”

GTAs Perceived as Mentors Help Resolve Group Challenges

We found most students who perceived their GTA as a mentor or a combination of mentor and instructor did not experience as many group challenges as students who perceived their GTA as instructor or other combinations. Additionally, most students who perceived their GTA as at least a mentor also reported their GTA to help resolve their group challenges. From our findings, students who perceived their GTA to help them resolve their issues also considered their GTA to be a mentor. It is possible that if a student perceived their GTA as a mentor their GTA more than likely helped students to create a safe space to work together. Furthermore, being relatable and approachable, both descriptions of a mentor by students in this study, have been reported to create a supportive classroom environment which can positively impact student buy-in for student-centered activities like group work [129].

Additionally, we found most students who perceived their GTA as a mentor or mentor and instructor were enrolled in Physics II mini-studio sections. It is possible students might have different expectations of how their GTA should interact with them if they already completed the first part of the sequence. Some students enrolled in Physics II mini-studio sections also described experiencing a situation similar to the scenarios during Physics I mini-studio.
Implications for GTA Preparation

*GTAs might recognize group challenges but choose not to engage*

In the teaching intervention (described in Chapter Four), we made the challenges experienced by each (avatar-)student group to be transparent, but in the classroom the challenges are not always obvious, as mentioned by Student F2 in an earlier section. It is also possible GTAs might not feel comfortable with engaging with groups about their group dysfunction or know what to do during the interaction. For example, during the group management similar session, Students E1 and E2’s GTA said in their reflection, “For the groups like CJ and Ed, I don’t, I then [short pause] because it is difficult for a TA to understand who is doing what with work, like in the second group.”

While GTAs might have difficulty recognizing or helping students resolve their issues, it is important for GTAs to help students with their group dysfunctions. Student B1 supports this claim when they wanted their GTA to be more communicative with students, “But as a TA itself, I like if you’re going to be a TA I feel like your job is to come in to teach and like be able to communicate with people and I felt like she was also very shy in that end and more quiet tone.”

While we found GTAs to help students resolve content-related group challenges only one student report their GTA to help resolve a group dynamics issue, despite eight students reporting to have experience such situations in their mini-studio section. Thus, GTAs might need training focused on identifying when groups are not functioning by putting aside the physics content. In their book chapter, Fonteijn and Dolmans provide examples of different group dysfunctions and suggestions of how to facilitate productive
group discussions [130]. Future microteaching or simulator sessions can give more explicit examples related to putting the content aside like checking in with student groups about their issues other than content or encouraging other students in the group to share their idea. Nguyen et al. also recommend strategies like encouraging student groups to create group policies and implement group roles to hold students accountable for their contribution to the group [129]. Future iterations of the teaching intervention will include discussions to help GTAs pinpoint less transparent student challenges in the classroom and clear strategies to use when managing group dynamics similar to the strategies previously described.

Conclusion

In summary, we found student enrolled in a mini-studio section to experience group challenges as described by the scenario prompts; Scenario One was the most reported scenario prompt to be descriptive of student group challenges. We also found students reported two other group challenges related to physics content: “all group members feel stuck” and “disagreed on how to solve problem”.

While most students who reported their GTA to not recognize their group challenge described by the scenario prompts, most students who reported an alternative group challenge related to physics content recognized their group challenge and helped to resolve their group’s issue. Additionally, most students who described their GTA to be a mentor or combination of mentor and instructor reported less group challenges or their GTA recognized and helped to resolve their group challenge.
Limitations

It is possible GTAs were using group management strategies more frequently than we interpreted here. We are not attempting to quantify GTA’s use of group management strategies or their awareness of group challenges, rather we are exploring examples of group management strategies used by GTAs perceived by students as well as which types of group challenges students reported their GTA helped resolve.

In addition, we rely on the perspective of students rather than observed accounts of GTA behavior in the classroom. It is possible by interpreting the student perspective of group management, our findings are biased towards student satisfaction. An example of this is Student E2.

We recognize the make-up of the student group might contribute to the existence or absence of group challenges. For example, when investigating small group work in biology, Theobald et al. found “working with a friend” was a predictor of student comfort in their group and their performativity [131]. In this study, several students described their group to be “friendly” or a “friend group”, or they had a friend in their group which might have impacted if they reported a group challenge since they are comfortable with their group members.

Future Work

To address the limitations discussed, future work will investigate the GTA perspective of group work and group dynamics by conducting interviews with GTAs. GTA and student perspectives could be compared, especially if GTAs were asked about
the same scenarios provided to the students during interviews. Additionally, GTAs’
comfortability with group work and group dynamics will be explored.

In future work, student group make-up will be investigated alongside students’
perspective of group work.
CHAPTER EIGHT: STUDY THREE – IMPACT OF TARGET SKILLS ON STUDENT FEELINGS

Introduction and Background

The use of evidence-based practices by STEM instructors supports student-centered instruction. For example, cold call can be used by an instructor to create the norm that student participation is expected during class. Dallimore, Hertenstein, and Platt found when instructors frequently use cold call, more students will volunteer their ideas throughout the semester [39]. However, it is possible that some students might not feel comfortable with sharing their ideas because of their fear of negative evaluation. Cooper, Downing, and Brownell found fear of negative evaluation to be at the root of the increase in student anxiety associated with cold call [40]. An instructor might choose to use error framing paired with cold call or other student idea probing techniques, like Stretch-It (questions that “stretch” the discussion beyond a student’s initial idea [96]). Error framing is an apprehension reduction technique that can be used to curb students’ fear of negative evaluation when sharing an idea [46,99,100].

Research about STEM GTAs’ use cold call, error framing, and Stretch-It, including how students feel about such skills in a small physics classroom, is limited. Since we expect physics GTAs to use each skill during their rehearsal sessions and in the classroom, and student feelings influence motivation for learning [128], we find it necessary to investigate how students feel when their GTA used the target skills in their mini-studio section, as informed by Reeves et al. about the impact of GTA training outcomes on students [27]. Additionally, the results of this study will be integrated into
discussions held by facilitators during future iterations of the teaching intervention described in Chapter Four.

Research Questions

To investigate student feelings about their GTA’s use of the target skills in their mini-studio section we ask the following research questions:

1. Do students perceive their GTA to use the target skills from the simulator training in their classroom or lab?

2. How do students feel when their GTA uses the target skills in the classroom or lab?

Additionally, we investigate ways GTAs responded to incorrect student ideas and how students felt as a comparison to their feelings about error framing. We ask:

3. What other ways do students report their GTA to respond to incorrect ideas, and how do they feel about them?

Methodology

Using Case Studies to Explore Student Perspective of Pedagogical Skills

Following a multiple-case study methodology [105,106], we explored how students feel when their GTA used the target pedagogical skills from the simulator training. We considered each student to be a single case since students have different perspectives of the atmosphere in their mini-studio section and GTA, even if some students had the same GTA. During analysis, we compared each student’s experience
or case and discuss common themes in our Findings and Discussion section like a case study by Sandi-Urena and Gatlin [64].

*Frameworks Utilized*

In this study, we investigate the influence of GTA use of target skills on student feelings, a relationship outlined in Reeves et al. [27]. By investigating this relationship, we can modify the teaching intervention to better support GTAs’ development of evidence-based strategies. Fear of negative evaluation (FNE) describes the negative feelings students might feel when their GTA uses a strategy like cold call and increases a student’s feelings of anxiousness [40]. We operationalize the actions GTAs take to mitigate student’s feelings of anxiousness or elicit positive feelings as protective facework [41,42]; actions taken by GTAs that elicit negative feelings are considered harmful facework [41,42].

*Participants*

Fourteen out of fifteen participants in Study Two were included in this study; one student (Student B1 from Study Two) did not respond to questions about their GTA’s use of the target skills from the simulator training and will not be included in this study. Each student was enrolled in either a Physics I or II mini-studio section participated in a semi-structured interview about their interactions with their GTA and their perception of their GTA’s use of the target skills. The level of experience teaching the mini-studio course and participation in simulator training varies across GTAs. Table 9 displays the enrolled course and GTA teaching experience for each participating student.
Table 9: Undergraduate Student Participants

<table>
<thead>
<tr>
<th>Student</th>
<th>Enrolled in Physics I or II</th>
<th>GTA's Level of Experience Teaching Mini-Studio Course</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>I</td>
<td>New/ Taught other</td>
</tr>
<tr>
<td>B*</td>
<td>I</td>
<td>New</td>
</tr>
<tr>
<td>C1</td>
<td>I</td>
<td>Returning</td>
</tr>
<tr>
<td>C2</td>
<td>I</td>
<td>Returning</td>
</tr>
<tr>
<td>D</td>
<td>II</td>
<td>New/ Taught other</td>
</tr>
<tr>
<td>E1</td>
<td>II</td>
<td>New</td>
</tr>
<tr>
<td>E2</td>
<td>II</td>
<td>New</td>
</tr>
<tr>
<td>F1</td>
<td>II</td>
<td>Returning</td>
</tr>
<tr>
<td>F2</td>
<td>II</td>
<td>Returning</td>
</tr>
<tr>
<td>G1</td>
<td>II</td>
<td>Returning</td>
</tr>
<tr>
<td>G2</td>
<td>II</td>
<td>Returning</td>
</tr>
<tr>
<td>H1</td>
<td>II</td>
<td>Returning</td>
</tr>
<tr>
<td>H2</td>
<td>II</td>
<td>Returning</td>
</tr>
<tr>
<td>I</td>
<td>II</td>
<td>Returning</td>
</tr>
</tbody>
</table>

*Student B is Student B2 in Study Two.

Undergraduate Student Interviews

Students who participated in the student interviews conducted at the end of Spring 2020 (data analyzed in Study Two) responded to questions about their GTA's use of cold call, error framing, and Stretch-It, and how they felt when their GTA used these skills. Three researchers (T.W., A.G.G. and C.M.D.) interviewed fourteen undergraduate students enrolled in either a Physics I or II mini-studio section taught by a GTA who participated in the simulator training. Each student individually participated in a semi-structured interview about their interactions with their GTA and peers in their mini-studio section. Due to the COVID-19 pandemic, each interview was remotely conducted using video-conferencing software (i.e., Zoom) to ensure the safety of
participants and interviewers [126]. The interviews were audio-recorded and transcribed.

We did not explicitly ask students if their used GTA cold call, error framing, or Stretch-It because students might not be familiar with that terminology. For cold call, we asked students if their GTA called on non-volunteering students to share an idea. For error framing we asked students to provide an example of how their GTA would respond to incorrect or unexpected student ideas. Since Stretch-It is a specific questioning strategy, we asked students about their feelings about their GTA’s use of questioning in general. The student interview protocol is available in Appendix E.

Data Analysis

Investigating Inter-Rater Reliability

Similar to Study Two, two researchers (C.M.D. and J.C.C.) investigated IRR for each code from the code book in two stages. During stage one, both researchers used the code book to code two (out of 14) interviews. After individually coding the two interviews, both researchers compared coding and discussed until agreement. Minor changes to the code book were made before the next stage of coding. In stage two, the same researchers individually applied the code book to the rest of the interviews (thirteen). Gwet’s AC1 was used to calculate a value that can be interpreted for investigating IRR [119]. The grain size for each code category was a student interview. Gwet’s AC1 values ranged from 0.73 to 1 across codes before discussion. After discussion, Gwet’s AC1 values for all codes were 1. A value of greater than 0.61 can be
interpreted as moderate agreement while a value greater than 0.81 can be interpreted as near perfect agreement [120].

Findings and Discussion

Cold Call

In our analysis, we find half (seven) of the students reported their GTA to use cold call in their mini-studio section. Even if a student reported their GTA to use or not use cold call in their mini-studio section, students might not feel comfortable with the strategy in general. During our investigation, we discovered two main feelings students associated with cold call: feeling nervous or feeling comfortable. When feeling nervous, students reported the nervousness to be either conditional or continuous.

Some Students Feel Nervous When Called On

One student ascribed their nervousness when called on to not having a correct response when answering a question, a concept linked with fear of negative evaluation. Student C2 explained “Whenever I do get called out in the small group, or in any other class, I feel like a little bit of pressure to get it right and get it right explaining it as well”. From their perspective, when their GTA used cold call in their mini-studio section Student C2 felt they had to provide a correct answer with correct reasoning in order for their response to be a valuable contribution.

Even though they did not report their GTA to use cold call in their mini-studio section, Student F2 and Student G2 both expressed a feeling of nervousness if their GTA were to use that strategy. Student F2 explained that when they have been called
on in other courses, they had a physical reaction. They said, “Yeah, I started stuttering. And I get really red. I really don't like [being called on]”. Student G2 expressed their nervousness associated with cold call might be due to a general anxiousness of speaking in front of a class. They explain, “Um, for me [pause] I'm kind of like, I'm kind of like, more nervous, like, class speaking”. From these findings, it can be implied that the use of cold call might increase a student’s feeling of anxiousness in the classroom.

Student Nervousness Associated with Cold Call Might Be Conditional

When discussing how they felt about cold call, five students described a conditional feeling of nervousness. Student D, Student E1, and Student G1 reported their GTA to use cold call in their mini-studio sections. When discussing how they felt about cold call, they all described a conditional feeling of nervousness if they did not know the answer. For example, Student G1 explained, “it really depends on like, what, what he’s asking me to solve, like, and how comfortable I feel with it. If it's something that I know and like, have worked on before, and I feel like I could solve it well, then like, I feel okay doing that. But if it's something new that I'm not really familiar with, or haven't really gotten to yet, I might feel kind of like nervous or not really or not prepared to do it”. This implies that Student G1 might have felt comfortable answering a cold call question if they were prepared to answer. Student E1 discussed a similar feeling. They expressed, “Um, it might have, it might have made me a little anxious if I didn't know really like what was going on, on. And I feel like sometimes, it was hard to know, like, some of the concepts since we were ahead of lecture”. In Student E1’s response they discussed a feeling of anxiousness with cold call because they often did not know the
physics content. Thus, students might feel nervous with cold call only when they do not feel prepared to answer.

Student D felt similar to Student G1 and E1, but they became comfortable with their GTA’s use of cold call over time. Student D exclaimed, “At the beginning, it was a little bit more nerve wracking, trying to, like, explain stuff in front of the class. But then after a while, you get used to everybody.” It is possible their GTA’s use of cold call early in the semester helped shape the expectations for the course. This is evident when Student D continued their thought about cold call by stating, “But it wasn't like you were put on the spot. And if you if you don't know it, you're gonna fail the class. It was nothing like that. It was like, you know, it was a very supportive environment.” Thus, the nervous feeling associated with cold call might diminish if GTAs create an environment that supports students sharing of ideas.

Even though they did not report their GTAs to use cold call, Student A, Student E2, and Student I felt a conditional feeling of nervousness if their GTA were to use the strategy. Student I exclaimed, “It depends if I know the answer or not. If I don't, then [pause] awkward.” In Student I’s case, not know how to answer a cold calling question might cause an awkward social situation in the classroom. Student E2 described what that situation might be in their mini-studio section. They expressed, “I would say like, it's uh depends [on] if I know the material or not. So, if I don't know the material, and I'm called on, then I'll feel, I'll feel very hesitant or ashamed that I don't know. Cuz it's like, well, I don't want to be wrong. And I don't want to look stupid in front of my peers.” Here, it is implied Student E2’s nervousness with cold call is due to fear of negative
evaluation, like Student C2. Thus, again, fear of negative evaluation causes a student to feel nervous with cold call.

Student A also felt a conditional nervous feeling associated with cold call if their GTA were to use such a strategy. They stated, “I personally do not like [being called on]. Because if I'm not confident in the answer, then I don't really know how to answer the question. Unless I'm confident, yes, then I'm fine. But in this case, I was not always prepared.” Student A’s mention of “not always prepared” is similar to Student E1’s “hard to know” indicating some students might be nervous in general if put on the spot to answer a question about physics. Thus, it is important to pair anxiety causing teaching strategies with other strategies focused on making student mistakes a positive learning experience.

Some GTAs Created a Comfortable Environment for Using Cold Call

Two students felt comfortable with the use of cold call in their class. Student C1 reported their GTA to use cold call in their mini-studio section and described their feelings when their GTA used cold call, “Well, I never felt uncomfortable. The class was very nice. And everyone was very respectful. And no one no one was disrespectful. So, I felt pretty comfortable and, and in my class.” However, Student F1 did not report their GTA to used cold call in their mini-studio section, but they expressed they would feel comfortable if their GTA were to use the strategy. They said, “Like I said, I mean, I, I personally didn't, it felt very low pressure, like I didn't feel tense or nervous going in there. It was pretty laid back so I wouldn't feel much of anything really.” This finding
implies that students might feel comfortable when their instructor uses cold call if the classroom environment is low-risk and relaxed.

Setting Expectations Early in the Semester is Crucial for Student Comfort with Cold Call

When asked how frequent their GTA used cold call Student C2 responded, “Before spring break, he did it for like the last two labs.” Since students were expected to attend thirteen mini-studio sessions and Spring Break occurred towards the end of the semester, it is possible the GTA challenged the students’ expectation of classroom norms when insisting on non-volunteering students sharing an idea in front of the class. If the GTA used cold call towards the beginning of the semester, like Student D’s GTA, then students would get used to the expectation of sharing ideas and the nervousness of answering a cold calling question might diminish.

GTA Responses to Incorrect Student Ideas

To investigate whether students perceived their GTA to use error framing, we asked students about their GTA’s responses to incorrect or unexpected student ideas. Additionally, we asked how they felt about their GTA’s response to compare their feelings associated with their GTA’s use of error framing with other ways GTAs responded to incorrect student ideas. Here, we present how the participants perceived their GTA to respond to incorrect ideas and how they felt about each response. We also make connections between the GTA responses and facework.

We found six different ways a GTA might have responded to an incorrect student idea, including error framing. Table 10 displays the six categories including descriptions
and example quotes from participants. The categories are not mutually exclusive. For example, a GTA might have used an error positive statement and provided an explanation.

**Students Felt Comfortable or Supported with Error Framing**

All three students who reported their GTA to use error framing also described a positive feeling about their GTA’s response to incorrect student ideas. Student E1 and Student H1 expressed their anxiousness was alleviated or they felt comfortable when their GTA used error framing. The interviewer (C.M.D) asked Student E1 about their GTA’s use of error framing made them feel, “Um, and so earlier, you mentioned like, let's say a student gave an incorrect answer. You mentioned that your TA would probably say something like, you know, ‘I can see where you're coming from’, or something along those lines. So, like, how did that make you feel knowing that if he were to respond that way, and you answered incorrectly?” Student E1 then responded, “I think it brought, like a feeling of like, comfortness, you know, like, even if I answered incorrectly, I know, that I won't be looked at, um in a way that like, I don't know, what I'm talking about, or like, I won't be thought of as like, ‘Oh, she's stupid’, you know, um, it just, it made it a very relaxed environment where it was easier to really learn.” It is evident in Student E1’s response that their GTA’s use of error framing mitigated their fear of negative evaluation when sharing an incorrect idea. Student H1 also felt comfortable sharing ideas when their GTA used error framing. When asked if their GTA’s response might help them feel less “unnerving” with cold call, Student H1 agreed by saying, “Oh, yeah, very much. So it was, you know, very approachable.”
Table 10: GTA Responses to Incorrect Student Idea Descriptions

<table>
<thead>
<tr>
<th>GTA Response</th>
<th>Description</th>
<th>Example Quote</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error Framing</td>
<td>GTA frames student mistake as natural or beneficial to the learning process</td>
<td>Student D: “He was just like, &quot;Okay, I see where you come from, like, I see, like, how you tackled that and good job and like trying stuff&quot;.”</td>
</tr>
<tr>
<td>Error Positive Statement</td>
<td>GTA praises student mistake or the effort that led to the mistake</td>
<td>Student G1: “…he would just encourage us to share ideas and be like, 'Well, it doesn't matter. Like, doesn't matter if you're right or wrong. Like, I just want to hear what you guys have to say and everything, like it would help if you participate and all that stuff'.”</td>
</tr>
<tr>
<td>Ask for Another Student</td>
<td>GTA asks for another student to answer or for student to provide another answer</td>
<td>Student I: “She would either say, “that's not exactly it” or “can you think of something else?” or just kind of look around the room, kind of like showing you that was not the right answer. And if anybody else wanted to give their input and they could raise her hand and say what they thought was the right answer.”</td>
</tr>
<tr>
<td>Response</td>
<td>Provide Explanation</td>
<td>Student F1: “If there was something put, like she was showing us an example of a problem on the board and then someone, and she said, ‘Oh, how do you do this?’ and someone said something wrong or whatever, someone was confused about something, she would just directly address it and show, like, how this would be done, or why it worked that way.”</td>
</tr>
<tr>
<td>Ask for Work or Reasoning</td>
<td>GTA asks students to describe their work or provide reasoning behind their answer</td>
<td>Student F2: “Um, she would direct this directly back to say, the homework or the quiz something with the similar concept. And she’s like, ‘Oh, so you did this? How? How did you get there?’ And you would explain to her like, okay, so following that logic…”</td>
</tr>
<tr>
<td>Ignore Student Idea</td>
<td>GTA does not acknowledge the student's idea, might be described as &quot;brushed off&quot;</td>
<td>Student B2: “She would kind of just brush it off and ask somebody else. She wouldn’t explain to us why we were wrong.”</td>
</tr>
</tbody>
</table>
Student D described feeling a different positive feeling than Student E1 and Student H1. They felt supported by their GTA's response to incorrect student ideas such that they felt comfortable with answering a cold calling question posed by their GTA. Student D expressed, “Recently, I have not had any professors that will make you feel bad about getting something wrong, or TA like that, but in the past I have. And so that kind of keeps you from wanting to answer, or like if you’d like for, like voluntarily or like it kind of makes you shy if you get picked on to like answer like un-voluntarily. But it wasn’t like that, because [GTA]'s like, real cool. He didn’t make you like-, if you got something wrong, it wasn’t like the end of the world. He was just like, "Okay, I see where you come from, like, I see, like, how you tackled that and good job and like trying stuff.” These findings imply that the use of error framing can mitigate negative feelings with cold call and create an overall positive learning environment.

All three student cases described error framing to elicit positive student feelings because they did not appear “stupid” when they shared their idea, or it was expected for students to share incorrect ideas. The above examples are similar to “work to avoid making you look bad” described by Gaffney and Gaffney as protective facework [42]. Thus, error framing is an example of protective facework.

Students Felt Supported by Other GTA Responses

Students also felt supported when their GTA used error positive statements combined with another strategy. Student H2 mentioned their GTA combined error positive statements with ask for another student response, “It's definitely, um, it's better than just hearing someone say no, because once you hear no, I don’t know, if, if the
person is actually gonna listen after that, or they're just like, no. And they kind of shut down. So, it's a good thing of like, okay, we're kind of there. Now, let's get all the way there. So. Yeah, you never feel like you're lost. At least to me. It was just a good okay. Listen up. So, what the full answer is or what I'm missing.” Student G1 described their GTA to pair error positive statements with providing explanations, “Um, I'd say just encouraged and motivated to work, you know, and I didn't always, like, share my ideas, but if I did, then I knew that it was okay, you know, if it wasn't right, you know.” From these findings, when students felt supported by their GTA’s response to incorrect ideas, they described their GTA to care about their learning, another example of protective facework [42].

However, it is difficult to disentangle whether the use of error positive statements is the only reason why students might feel supported. For example, Student C1 felt supported by their GTA who only provided explanations when responding to incorrect student ideas. “Oh, I mean, I felt supported. Like, I, I felt good. And I felt like I could continue on with the problem, and then I would do it correctly.” It is possible that students expect their GTA to provide explanations when learning physics and feel dissatisfied when their GTA does not do that [124].

Some GTA Responses Elicited Student Feelings of Fine or Neutral

Unlike Student G1 and Student H2, Student A reported feeling neutral with regards to their GTA’s use of error positive statements when responding to incorrect student ideas. They said, “It made me feel like, even though I didn't have the answer he was looking for. I still had, I still was, had the basic knowledge to at least figure it out.
So, I didn’t feel that bad.” Student G2 and Student I also felt neutral about their GTA’s use of ask for another student response. Student G2 said, “Um, normal? Fine. I guess.” and Student I said, “Nothing, it was just kind of, they got it wrong. And that’s pretty much it, and we move on.” Student C2 also felt neutral when their GTA asked for work or reasoning when they said, “Kind of neutral.” Here, some students, like Student A and Student G2, described a positive inclination of feeling neutral about their GTA’s response. Both student responses are similar to facework examples “work to avoid making you look bad” and “make sure s/he doesn’t cast you in a negative light” [42].

Some Students Felt Rushed or Frustrated

Student F2 described their GTA to respond to incorrect student ideas by providing explanations and asking for work or reasoning. They expressed their feelings with this strategy, “Sometimes you were a little frustrating, because I feel like you, you expect a really straight answer. But at the same time, they were a lot more helpful than her just telling me what the answer was. I felt like I really understood after because I had to think of the solution”. While Student F2 felt frustrated with their GTA asking questions, they also found the questions useful for their learning. It is possible students might feel frustrated when learning a new subject and give up. Student F2’s GTA expressed experiencing student frustration with their use of Stretch-It in the classroom in Fall 2019 which might have contributed to their abandonment of the skill, as part of the discussion in Study One. It is possible if a student shows frustration towards the use of questioning that their GTA will stop asking questions and start explaining.
However, Student B2 only described negative feelings with their GTA’s response. Student B2 mentioned their GTA “brushes off” student ideas and asks for another student response, making them feel rushed as they explained, “I feel like maybe she was just rushing it just to try and get the right answer. So, we can continue on with our lab. You know, rather than caring about, okay, you’re close, but you’re forgetting something, you know, something like that, or working with that person rather than going to ask somebody else.” While asking for another response might make a student feel neutral as shown in Student G2 and Student I’s responses, ignoring a student’s idea might be considered harmful facework according to Student B2’s response of their GTA not caring.

Most Students Felt Comfortable with Their GTA’s Use of Questioning

Since Stretch-It is a specific questioning skill and we were concerned students might not be familiar with such skill, we asked students about their GTA’s general use of questioning. Here, we focus on if students felt interrogated by their GTA’s use of questioning.

We found only one student (out of fourteen students) felt interrogated by their GTA’s use of questioning. When asked if they felt interrogated by their GTA’s use of questions Student B2 expressed their feelings, “Yeah, I definitely got those. And my group would definitely agree with that, because we’d talk after and be like [whelp or oop noise] so what’s the answer?” It is possible the GTA’s use of questioning caused confusion because the student did not know what the answer to the initial question was before the GTA asked another question. While we do not discuss classroom
observation data in this study, we observed Student B2’s GTA to ask the highest percentage of questions in any setting (whole class or one-on-one), namely 51.5% of two-minute observation intervals on average across their Spring 2020 observations (38.3% was the second highest percentage of questions asked by Student I’s GTA). When asked if the student felt annoyed or stressed by their GTA’s use of questioning, Student B2 responded, “Yeah, I definitely got that as well. If anything was more like, stressed out, and like, confused, a little more confused than I already was.” It is possible when a GTA asks a high percentage of questions without pausing for students to respond or think about the answer a student’s feeling of anxiousness might increase. This finding is comparable to “make it hard for you to purpose your own ideas” which is considered harmful facework [42].

**Conclusion**

In summary, we found most students felt nervous when their GTA used or might use cold call in their mini-studio section. Additionally, some students attributed a condition to feeling nervous, especially if they did not know the answer to the cold calling question. This is consistent with other findings that students described an increase in their feelings of anxiousness when their instructor used cold call [39,40]. However, two students reported to feel comfortable sharing their ideas if their GTA called on them in their mini-studio section.

Additionally, we reported six different ways GTAs responded to incorrect or unexpected student ideas. Error framing was reported by three students to be used by their GTA in response to incorrect student ideas and each student felt their anxiousness
to decrease or felt supported by this type of response. One student expressed they would feel comfortable answering a cold calling question when their GTA used error framing. These findings indicate that error framing is an act of protective facework and mitigates the negative impact of negative evaluation on student comfort with sharing their ideas.

We also described a mixture of positive and neutral feelings for error positive statements, ask for another student response and provide explanation. Some students felt supported when their GTA used error positive statements with another type of response. Other students reported a neutral feeling when their GTA only used error positive statements, ask for another student response, or ask for work/reasoning. Some students who felt neutral also provided positive inclinations of their GTA's response which were like examples of protective facework.

We found two students to report negative feelings associated with their GTA's response to incorrect student ideas. One student felt their GTA's use of explanation and asking for work/reasoning was frustrating but helpful to their learning. We also reported one student to feel rushed when their GTA ignored their incorrect idea because they felt like their GTA did not care about their learning, opposite of the protective facework example, "show that he/she cares about your learning experience" [42]. This implies that when a GTA brushes off a student's idea, the GTA is performing harmful facework.

Finally, we reported all students except one did not feel interrogated by their GTA's use of questioning. From this finding, a GTA's use of questioning might be
interpreted to be harmful facework when GTAs do not give students opportunities to think or respond to questions.

Limitations

It is possible students who were already comfortable with sharing their ideas participated in the interviews, and perspectives of those who might have social anxiety or communication apprehension are not reflected in our findings. We also did not ask students if they have general anxiety or depression which might have affected their feeling of anxiousness with cold call or other strategies mentioned in this study, as described in [40].

Another limitation is the size of our sample. Since we only interviewed one or two students per GTA and they might be enrolled in different sections, it is possible we did not reach saturation of the overall environment created by each GTA. Also, we did not include the GTA perspective of the use of the skills or classroom observation data. Therefore, the results of this study are limited to the perspective of students. It is possible GTAs might recognize the negative impact of cold call on student feelings but are not sure how to mitigate the impact.

Future Work

Future work will address the limitation of students with varying levels of anxiety in the classroom by including interviews questions or a questionnaire similar to the one used in [40]. We also plan to interview more students. We will use the results from those
interviews and this study to create a survey to distribute to students enrolled mini-studio sections.

For comparison to the results of this study, we plan to interview GTAs about their perspective of their skill use, the impact of the skill on their students’ feeling, and strategies they use to mitigate feelings of anxiousness caused by FNE. Additionally, it is possible GTAs might have conceptualized error framing in a different way than the training facilitators. We document examples of chemistry GTAs conceptualizing in a different study [34] and plan to continue our investigation with physics GTAs.

In this study, we discussed a GTA’s use of questioning in general rather than the specific types of questions presented in Study One. Future work we will explore students’ perceptions of their GTA’s use of Stretch-It. To explore their perspective of Stretch-It, we will modify the interview protocol to include examples of Stretch-It used by their GTA documented during classroom observations or simulator sessions and ask students about their feelings towards each example.

In addition to the topics included in this study and Study Two, we also presented students with examples of facework and asked them to discuss if their GTA used each item and examples of what their GTA did. We plan to analyze this portion of the interview for both Fall 2019 and Spring 2020 cohorts.
CHAPTER NINE: CONCLUSION

This dissertation investigated the impact of using a mixed-reality simulator for physics GTA training on GTAs and students. In this chapter findings and implications from Study One through Study Three will be discussed to answer the overarching research question: What is the impact of teaching rehearsal in a mixed-reality classroom simulator on physics GTAs and their students in active learning lab and recitation sessions?

Study One – Impact of High-Intensity Training with a Mixed-Reality Simulator on GTAs’ Use of Questioning

This study focused on investigating the impact of the simulator training on GTAs’ use of the Stretch-It target skill by answering the following questions:

1. Do GTAs change their use of questioning in simulator sessions after rehearsing the Stretch-It skill?

2. What is the impact of practice during simulator sessions on GTAs’ use of questioning in their actual classroom?

To answer the first question, we explored the high-intensity training (Fall 2019) cohort’s use of Stretch-It during each rehearsal session because their participation in the training is a moderating factor for transfer of the target skill use to the classroom [26].

In summary, we found strong evidence that skill rehearsal paired with facilitator feedback directed at GTAs’ use of Stretch-It during their sessions with the simulator
influenced GTAs’ practice of Stretch-It during rehearsal sessions. Namely, we observed GTAs to increase or shift their use of Stretch-It across rounds during the session focused on using Stretch-It. Additionally, we observed all four sub-categories of Stretch-It to be used by GTAs during their practice in the simulator sessions when facilitators provided focused feedback about using Stretch-It, sessions two and four. This finding implies that a deliberate practice model with specific facilitator feedback is an important factor for GTA engagement in training.

To answer the second question, we evaluated whether GTAs’ use of Stretch-It questions in their mini-studio sections changed after the simulator session on Stretch-It. We compared classroom observation data for two GTA cohorts, pre-simulator training (Fall 2018) and high-intensity simulator training (Fall 2019) semesters.

Here, we found an immediate impact of the simulator training on GTAs’ use of Stretch-It: Explain Logic during the high-intensity simulator cohort semester, but their use of the skill dropped back down to pre-simulator training averages throughout the semester. Specifically, we observed more GTAs to use a higher rate of Explain Logic during the classroom observation immediately following practice of Stretch-It in the simulator. We also observed more GTAs in the high-intensity training semester to use a higher rate of Explain Logic than GTAs in the pre-simulator semester which suggests more GTAs are intentionally using Explain Logic in the classroom after the training. However, the simulator training did not impact GTAs’ use of Follow-Up in the classroom; GTAs might randomly use Follow-Up because of the curriculum or other factors such as GTA characteristics or cognition. We find that GTAs may shift their practice after
training in the simulator combined with facilitator feedback, but the impact varies across the different categories of Stretch-It. Additional research is needed to investigate which skills can be impacted by the current training model and how to adjust the model to impact other skills.

**Study Two – Student Perceptions of GTA Strategies to Manage Group Challenges**

In this study we explored the similarity between our group work scenarios and student experiences as well as students’ perception of how their GTA helps to resolve their group dynamic challenges. We focused on answering the overarching research question by answering four sub-questions. Thus, we ask:

1. How similar are the group challenge scenarios created by the research team to group challenges experienced by students in the physics mini-studio section?
2. What group challenges do students experience that are not described by the group challenges scenarios?
3. How do students perceive their GTA to recognize and help resolve their group challenge?
4. How do students perceive their GTA’s role in the classroom and how is it related to their GTA resolving their group challenges?

We found students enrolled in a mini-studio section experienced group challenges similar to those described by the scenario prompts, especially Scenario One. In Scenario One, “Miguel” dominates the group’s decisions while “Sophia” remains quiet and seems frustrated. We also found students reported two other common group
challenges related to physics content: “all group members feel stuck” and “group members disagreed on how to solve problem”.

While most students reported their GTA did not recognize their group challenge described by the scenario prompts, most students who reported an alternative group challenge related to physics content reported their GTA recognized their group challenge and helped to resolve their group’s issue. Additionally, most students who described their GTA to be a mentor or combination of mentor and instructor reported less group challenges or their GTA recognized and helped to resolve their group challenge. We find GTAs recognize and help to resolve group challenges related to physics content but not group dynamics. Further investigation is necessary to explore why GTAs might not recognize or choose to ignore group dynamics issues. We also find GTAs who were perceived as a mentor were reported by students to help resolve their group challenges. More research is needed to explore training efforts that can support GTAs to put aside the content to focus on group dynamics.

Study Three – Impact of Target Skills on Student Feelings

Since students might feel uncomfortable with their GTA’s use of the cold call or questioning because of fear of negative evaluation, this study investigated the impact of GTA’s use of the target skills rehearsed in the simulator training on student feelings. Additionally, this study investigated student perception of GTA responses to incorrect student ideas, including error framing, as potential actions of facework. Specifically, the following research questions were asked:
1. Do students perceive their GTA to use the target skills from the simulator training in their classroom or lab?

2. How do students feel when their GTA used the target skills in the classroom or lab?

3. What other ways do students report their GTA responds to incorrect ideas, and how do they feel about them?

We found most students felt (or would feel) nervous if their GTA were to call on them in their mini-studio section, regardless of whether they reported their GTA used cold call. Additionally, most students who felt nervous with cold call also discussed a condition when feeling nervous, like in a situation where they did not know the answer to the cold call question. This finding is consistent with other studies investigating the impact of cold call on student feelings of anxiousness [40]. We also found two students to feel comfortable with answering a cold calling question and one student felt comfortable with their GTA’s use of cold call over time. This finding may suggest that GTAs can create an environment such that students feel comfortable with sharing their ideas in a non-voluntary way.

Additionally, we reported GTAs to respond to incorrect student ideas using six different strategies. Error framing was only associated with positive feelings (e.g., decreased anxiousness and supported) and was reported to mitigate nervousness caused by cold call. We also described a mixture of positive and neutral feelings for error positive statements, ask for another student response and provide explanation. Students reported a neutral feeling when their GTA only used error positive statements,
ask for another student response, or ask for work/reasoning. Most GTA responses to incorrect student ideas paralleled examples of protective facework as described by Gaffney and Gaffney [42]. Ignoring a student idea was associated with negative feelings (e.g., rushed) and is considered harmful facework.

**Limitations**

In this dissertation, we explored the impact of simulator training on GTAs and students in one context, (e.g., one institution, one discipline, one course-type) whereas Reeves et al. states it is vital to investigate the influence of contextual variables [27], including investigating similar training across institutions and disciplines (like in [65]). Since we do not have a comparison of the impact of training on GTAs and students in other contexts, the results of this dissertation are limited and might not be generalizable to all contexts.

Additionally, Carr and Kemmis criticized the use of technical action research because of the lack of agency participants have on their own training and teaching goals [103]. They continued to discuss how the process of teaching innovation might drop off due to external members forcing teaching expectations on teachers [103]. This implies some GTAs might not have fully engaged with the training as intended by training designers or wished to participate in the research because they did not have a say in their training.

Another limitation of this dissertation is our use of case study methodology to investigate the student perspective of group work and student feelings. Case study methodology has been criticized as too context-specific such that findings might not be
generalizable to other settings [105,106]. This implies the findings in Study Two and Study Three might not be applicable in other courses or disciplines despite the rich details described in student responses. We also only investigated the student perspective using this methodology which might create bias in the findings as the GTA perspective was not considered.

Additionally, we did not investigate how other pieces of training or people might have influenced GTA teaching behaviors in this dissertation. As novice teachers, GTAs are continuing to develop their teaching identity throughout their academic career. GTAs might have received mixed messages about how to teach from places other than the simulator training, like faculty lecture instructors [125] and their teaching colleagues [132].

Another limitation is that we did not explore the interaction between the simulator training (or classroom environment) and identity (e.g., citizenship, gender, ethnicity, etc.). We recognize that identity might have influenced the findings. For instance, international GTAs might have different needs than domestic GTAs (examples included in [133]).

Furthermore, not every graduate student is interested in teaching and might not agree with using student-centered instruction during recitation or lab. We noticed during classroom observations, some GTAs would not spend the entire time allotted to recitation on the tutorial worksheet. Some students during interviews expressed their GTA to verbally admit to their disagreement with the implementation of the tutorial. It is possible a lack of interest in teaching and student-centered instruction softened the
impact of the simulator training on GTAs’ instructional practices. We cannot disentangle the effects of GTA “buy-in” for the curricula or student-centered instruction from our findings.

**Future Work**

Here, we address future studies in response to the limitations presented in this dissertation.

To investigate the impact of the simulator training in other contexts, future research will explore using the simulator training model presented here to train physics GTAs teaching other courses and in other disciplines (like chemistry). Comparisons should be made between the results and implications of this dissertation to the results of these future studies.

In future work investigating GTA training, we encourage researchers to follow an action research methodology with GTAs as either participatory or emancipatory participants [103]. GTAs should have a more active role in their preparation by providing them with opportunities to impact training design by tailoring modules to individual GTA needs and including GTAs in classroom observation data collection and analysis.

In this dissertation, we followed a case study methodology and only investigated the student perspective for Study Two and Three. Future work should continue to use case study methodology when investigating student perspective of their feelings and experience with group work in the same context to compare results with those of this dissertation. The same methodology can be used to explore the GTA perspective and
triangulation between both sets of data can be analyzed to create a more holistic understanding of the classroom environment. In the future, we plan to interview GTAs about their perspective of the group challenges mentioned by students in Study Two and their perception of facework described in Study Three.

To address the mentioned concern of multi-tiered professional development, future work will focus on designing a more cohesive program such that training objectives align with each stakeholder’s vision (e.g., course designer, department, etc.). Additionally, future training programs design will consider the impact of identity on training design, GTA engagement during the training, and GTA and student outcomes. This includes collecting data focused on connecting GTA “buy-in” to the training and GTA outcomes.

Future work is necessary to investigate which pedagogical skills are impacted by training with the simulator. In Study One, we found an impact of the training on GTAs’ use of Explain Logic but not Follow-Up.

Since we only reported student feelings about their GTA’s general use of questioning, we continue our investigation presented in Study Three by modifying the interview protocol to include examples of Stretch-It used by their GTA documented during classroom observations or simulator sessions; we plan to ask students about their feelings towards each example.
Takeaways and Implications for GTA Training

The focus of this dissertation was to investigate the use of the mixed-reality classroom simulator, TeachLivETM, on GTAs and students. The studies presented in this dissertation demonstrated an association between GTAs' participation in the simulator training and GTAs' instructional behaviors, which influenced student experiences with group work and student feelings.

By investigating GTAs' use of questioning during rehearsal sessions and in their actual classrooms, we find rehearsal of questioning in the simulator positively impacted their use of questioning in their actual classrooms. We also speculate feedback to be crucial to improving their use of pedagogical skills during training which translated to the classroom. Standard models of STEM GTA preparation lack opportunities for GTAs to receive feedback about their teaching, but our finding suggests rehearsal paired with feedback is important to improving STEM instruction. STEM GTA training designers and facilitators should consider using deliberate practice models (e.g., rehearsal paired with feedback) to support GTAs' development of complex pedagogical skills.

Through our analysis of student interviews, we find GTAs are successful at recognizing and resolving student group challenges caused by physics content. However, students expressed their GTA failed to recognize their group dynamics issues (e.g., one student dominates). In the simulator training, we incorporated a rehearsal session focused on interacting with avatar-student groups with group dysfunctions such that avatar-students explicitly voiced their group’s issues. However, in our findings, group challenges experienced by actual students might not be as visible. We speculate
GTAs might have difficulty identifying group dynamic issues if they are not actively engaging with a student group. Moreover, GTAs might not feel comfortable addressing group dynamic issues or know how to resolve the challenges. We recommend STEM GTA training designers and facilitators to hold discussions with GTAs about group dynamics that might be happening in their classrooms, provide GTAs with practice identifying group dynamic issues using a deliberate practice model, and give GTAs clearly defined group management strategies.

Across studies, we find evidence of student frustration in the classroom when their GTA used instructional practices that might have violated their expectations of how their GTA should behave in a physics class (i.e., GTA 1 and GTA 13’s use of Stretch-It discussed in Study One; Student E2’s description of a “disengaged” GTA in Study Two; and Student F2’s use of questioning in Study Three). We recommend when training STEM GTAs in student-centered classrooms, training designers and facilitators should discuss student resistance with GTAs and give clear strategies on how to manage student resistance [124,127].

Additionally, by analyzing student interviews, we find students perceived GTAs to use cold call and error framing, target skills from the training, in the classroom. While most students expressed feeling nervous if they were called on by their GTA, we find examples of reduced anxiousness with cold call because their GTA created a supportive environment. We also find students to report error framing to mitigate negative feelings of anxiousness when sharing an incorrect idea. From our findings we suggest cold call to be paired with a communication apprehension strategy to create an
appropriate classroom environment for cold call because of the extensive literature [39,40,47,48], including our findings, documenting the negative impact on student feelings of anxiousness or nervousness when they share out. We recommend if STEM GTA training facilitators suggest GTAs to use cold call in the classroom they should 1) pair cold call with error framing, 2) encourage the use of cold call at the beginning of the semester when setting class norms, and 3) provide GTAs with a space to practice and receive feedback about the use of both skills together.

Overall, our findings support the use of the GTA Professional Development Evaluation and Design framework when investigating new or existing STEM GTA training programs. We encourage training facilitators and researchers to investigate the impact of STEM GTA training efforts on GTAs’ teaching practices and student outcomes. Since this framework provides a general guide of factors that influence training design and GTA and student outcomes, it is possible to pair this framework with other frameworks to create a rich study design, as described below.

By using a situated cognition training model, we provided GTAs with opportunities to put pedagogy into practice and receive feedback. From our findings, we noted feedback was influential to GTAs’ rehearsal of pedagogical skills and use of such skills in the classroom. However, feedback is often neglected in STEM GTA training efforts. We suggest to training designers to use a situated cognition model when considering new ways to train GTAs. We also call on STEM departments to implement GTA training that uses a situated cognition model to provide scaffolding for the use of evidence-based teaching practices.
Not highlighted by the Reeves et al. but equally important to investigate is the classroom environment created by STEM GTAs. Since student-centered instruction depends on social interaction there is room for student feelings to be impacted by interacting with their peers and instructor. We used fear of negative evaluation (FNE) as motivation to explore student perception of group work and student feelings of anxiousness in the classroom. From our findings, we noted that students might feel nervous or anxious during group work or when GTAs use strategies that require students to share an idea, as a result of FNE. We suggest researchers to continue to use the conceptualization of FNE to investigate student feelings of anxiousness in physics recitation and lab sections. Additionally, we encourage STEM instructors, not just GTAs, to consider the impact of their teaching behaviors on their students’ feelings.

In this dissertation, we used facework to operationalize GTA actions that influence the social environment of the classroom. By using this framework, we made the connection between GTA teaching practices and student comfortability in the classroom. Along with FNE, facework can be used by researchers to explore GTA actions that help students feel comfortable sharing their ideas in front of their peers. While facework can be described as harmful or protective, we propose focusing on using protective facework when investigating GTA actions in the classroom, especially when following a critical research methodology, like Action Research.
APPENDIX A: SIMULATOR TRAINING– GTA HAND-OUTS
Newton’s Laws Tutorial – Cold Call and Error Framing (Session One) Activity

Description of Student Activity

Teaching skill: Whole class cold call/normalizing errors

Physics topic: Force and motion

Student learning objectives:
- Recognize the relationship between force and motion
- Be able to apply Newton’s second law to physics scenarios
- Reconcile Newton’s second law with common sense

Scenario: Students are working in small groups on a worksheet. In the worksheet, students consider a scenario in which a child who has fallen down a well, is being rescued. The child is fastened to a rope and pulled upward at a constant speed. Students are discussing whether the upward force by the rope is greater than, less than, or equal to the downward gravitational force.

Activity: Students are working on the questions in the worksheet. The worksheet prompts students to answer the questions in two ways: (1) use intuition or common sense and (2) use Newton’s second law. The worksheet assumes students would arrive at inconsistent answers. It then guides students to reconcile Newton’s second law with their intuition.

In the simulator: The physics GTA is discussing the physics scenario with the students. The GTA intends to elicit student ideas and reasoning by asking questions. Then the GTA guides students to reconcile Newton’s second law with their intuitive thinking.

Key takeaways from the worksheet:
- A net force is *not* needed to maintain an object’s motion.
  - When the net force is zero, an object, which is initially moving, would move at a constant speed in the same direction.
  - When the net force is zero, an object, which is initially at rest, would stay at rest.
- A net force is needed to initiate or change an object’s motion.
  - A net force is linearly proportional to acceleration, which describes the change in velocity per unit time.
  - A net force is needed to change the velocity of an object.
Student Worksheet

Summary of section I of the worksheet (Note that section II builds on section I):
In the first section of the worksheet, students consider a scenario in which a child who has fallen down a well, is being rescued. The child is fastened to a rope and pulled upward at a constant speed. The free-body diagram is shown at right (Note that students are not given this). Students are discussing whether the upward force by the rope is greater than, less than, or equal to the downward gravitational force.

1. The worksheet first asks students to answer the question above using their intuition.
2. It then asks students to answer the same question using Newton’s second law.

Section II. Refining intuition to reconcile Newton’s laws with common sense

Most students have, or can at least sympathize with, the intuition that upward motion requires an upward force, in which case the upward rope force must “beat” the downward gravitational force to make the child move up. Can we reconcile that intuition with the Newtonian conclusion that the upward force merely equals the downward force?

A. (Work together) Consider the child, initially at rest, right when the rope first starts to pull him upward. During that initiation stage of the motion, is the upward force from the rope greater than, less than, or equal to 250 newtons (the child’s weight)?
   1. What does Newton’s Second Law say about this question? (Hint: Is the child accelerating during the initiation of the motion?)
   2. Does the Newtonian answer here agree with common sense?

B. (Work together) Now consider the child’s motion after the initiation stage of the motion once he is already moving.
   1. Intuitively, if the rope’s force remains larger than the child’s weight (like during the initiation stage), does the child continue speeding up, or does he slow down, or rise with constant speed? Briefly explain.
   2. Does Newton’s second law agree with your answer? Explain.
   3. Intuitively, if the rope force became smaller than the child’s weight, would the child speed up, slow down, or rise at steady speed? Briefly explain.
   5. Let’s tie this all together. It makes sense that, if the rope force remains greater than the gravitational force, the child keeps speeding up; and if the rope force becomes less than the gravitational force, the child slows down. By this line of intuitive reasoning, what happens to the child’s motion if the rope force equals the...
child’s weight, i.e., if the rope force “compromises” between being greater than and being less than the child’s weight? Explain.
6. Does Newton’s second law agree with your answer?

C. Think about the “intuition refinement” diagram below.

1. Which of those two refinements were you using in part B above?
2. Which of those two refinements agrees with Newton’s Second Law?
3. Which of those two refinements were you using back in part I., question B. and part I., question C.?

Wait for the class discussion before you proceed!
Pendulum Lab – Stretch-It (Session Two) Activity

Description of Student Activity

Teaching skill: Stretch-It

Lab activity: Pendulum for pros

Student learning objectives:
Note that the lab is NOT intended to reinforce concepts related to the period of a pendulum, but rather to improve ability to make decisions with data.
- Overall goal: Conduct an experiment to evaluate whether the angle of amplitude affects the period of a pendulum
- Decide what and how much data are to be gathered to produce reliable measurements.
- Propose and carry out follow-up investigations or revisions in light of the data and model, particularly to improve the reliability of the data.

Scenario: Students have already measured the period for a pendulum at 10 degrees and at 20 degrees, respectively. For each angle, they calculated the average and the standard uncertainty, as well as the t-score they used to make a judgement on whether or not the period depends on the angle.

In the simulator: TAs will be posing questions in small groups to help students interpret their initial data set, and come up with ways to improve the quality of their measurements.
I. Part I
Note: Since the simulator activity focuses on part II, we briefly summarize what students have done in part I.

- Students have measured one period (the time it takes for a full cycle) of a pendulum in two ways:
  - Starting from its highest position
  - Starting from its lowest position

For TA (not for students), here's a simulation in which you can make measurements.


- Students have made several trials and calculated average, as well as learned to quantify the reliability of their data using standard uncertainty.

II. Part II
Note: Students have completed sections A and B. They are working on section C in the simulator.

A. Initial investigation
Write down a plan for a high-precision measurement of the period of a pendulum at amplitudes of 10 degrees and 20 degrees. Include a clear description of how you will determine the uncertainty in your measurements. Use the earlier discussion and even the data collected to inform your decisions.

Carry out your plan to measure and compare the period of the pendulum at 10 and 20 degrees. What does the comparison mean or suggest?
B. Quantifying comparisons
Now that we have a statement about the reliability of a data set and high-precision measurements of the period of a pendulum, we want to determine whether the period of the pendulum is the same at 10 degrees and 20 degrees, or if one is systematically different from the other.

Work with your group to come up with a way to quantitatively compare the period of the pendulum at each amplitude. Focus on inventing a method to quantify how distinguishable the two data sets are. To get started, come up with a list of features of the data that are important to consider in making this comparison. Note: You may find the graphs at right useful for identifying critical features.

Then move on to inventing a quantitative index. Here are some guiding ideas:
1. A small index should imply that the data are not very different (indistinguishable) while a large index should imply that the data are very different (distinguishable).
2. Your method should work for many data sets, not just the two we're working with.

We will have a group discussion about everyone's inventions.

C. Revised and improved investigation
Based on your interpretation of your initial data set, write a plan for improving the quality of your measurements. Discuss your results and your plan with other groups. Feel free to modify your plan based on this discussion, recording in your notes any changes that you make. In your plan, include a short discussion about why you chose that method.

Perform your revised measurements and analysis. In addition to comparing the results at 10 and 20 degrees, evaluate whether your improved measurement plan led to improved measurements (e.g., are your uncertainties in your measurements smaller this time?)
Background material

Physics concepts and relationships:
- Period (of a pendulum): the time it takes the pendulum to complete one cycle. By convention, capital T is used to denote period.
- Period formula for small angle approximation: $T = \frac{2L}{g}$, where $L$ is the length of the pendulum and $g$ is the acceleration due to gravity.

Concepts related to measurements and statistics:
- **Average (or mean):** The average can be calculated by the sum of all the values divided by the total number of values.
- **Uncertainty:** uncertainty describes how precise a measurement is.
- **Standard deviation:** It tells us, on average, how far are the data points from the average. If we took another data point, it will likely fall within the standard deviation from the mean (mean +/- standard deviation). That means that the uncertainty on any single measurement is the standard deviation. (However, a single data measurement is less reliable than the average of a collection of measurements!)
- **Standard uncertainty in the mean:** It is the standard deviation divided by the square root of the number of measurements.
- **T'-score:** the t-score takes the difference between the two sets of measurements and divided by the combined uncertainty of the difference. As such, the t-score gives a quantitative measure of how different the two measurements are relative to their uncertainties.

**Interpretation of t-score**

FIG 3. Interpretations of and follow-up behaviors from a t'-score comparison between two measurements.
A video that explains t-score:  

Statistics formulae:

- Average/mean: Sum all the data, then divide by the total number of data points. 
  Suppose we have 10 data points (10 trials of measurement).
  \[ \bar{x} = \frac{x_1 + x_2 + \ldots + x_{10}}{10} \]
  More general:
  \[ \bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n} \]

- Standard deviation:
  For 10 data points: \[ \sqrt{\frac{(x_1-\bar{x})^2 + (x_2-\bar{x})^2 + \ldots + (x_{10}-\bar{x})^2}{10}} \]
  More general: \[ \sqrt{\frac{(x_1-\bar{x})^2 + (x_2-\bar{x})^2 + \ldots + (x_n-\bar{x})^2}{n}} \]

- Standard uncertainty in the mean: standard deviation divided by the square root of data points.
  For 10 data points: \[ \delta = \frac{\sqrt{(x_1-\bar{x})^2 + (x_2-\bar{x})^2 + \ldots + (x_{10}-\bar{x})^2}}{10} \]
  More general: \[ \delta = \frac{\sqrt{(x_1-\bar{x})^2 + (x_2-\bar{x})^2 + \ldots + (x_n-\bar{x})^2}}{n} \]

- t-score: \[ \frac{\bar{A} - \bar{B}}{\sqrt{\frac{\delta_A^2}{n_A} + \frac{\delta_B^2}{n_B}}} \] where \( \bar{A} \) is the average of data set A, and \( \delta_A \) is the standard uncertainty in the average \( \bar{A} \); and \( \bar{B} \) is the average of data set B, and \( \delta_B \) is the standard uncertainty in the average \( \bar{B} \).
DC Circuits Tutorial – Group Management (Session Three) Activity

Description of Student Activity

**Teaching skill:** Small group discussion; practice with dysfunctional student groups/group dynamics

**Physics topic:** Direct Current (DC) Circuits; Ohm’s Law

**Student learning objectives:**
- Investigate the relationship between current and voltage in a DC circuit (Ohm’s Law).
- Investigate different DC circuit configurations.

**Scenario:** This module takes place while students are working in small groups on their worksheet. The GTA has tasked the students to work through and discuss concepts from the worksheet in their groups.

**Activity:** Students are in their small groups working on a worksheet about light bulbs, current, and Ohm’s law. The students have been prompted by the GTA to answer the questions and discuss the concepts in the worksheet.

**In the simulator:** The students are working on and discussing the worksheet in their small groups. The physics GTA is tasked with investigating how the students are proceeding through the worksheet and making sure the groups are working well (each student is participating and contributing their ideas to the group; one student is not dominating the group; disagreements are being met and no one’s idea is being ignored). Each group has progressed through the worksheet at different paces and will need help from the GTA to work together.

**Key takeaways from the worksheet:**
- The current measured across the battery of a circuit depends on the configuration of the components in the circuit.
- There is a linear relationship between voltage and current measured across a component in a DC circuit.
- A battery is not a source of constant current in a direct current circuit but a source of constant voltage across its terminals.
III. Circuit with more than one bulb [3]

A. Work individually for this part. Imagine a two-bulb circuit with the bulbs connected one after the other as shown. Don’t set it up yet. Predict how the brightness of the two bulbs will compare to the brightness of an identical bulb in a single-bulb circuit. (Two bulbs connected one after the other are said to be in series.)

Predictions can vary.

B. Compare your reasoning with that of your partners. Again, you don’t have to agree with one another, but discuss your answers until you understand each other’s reasoning.

C. Now, as a group, gather the needed the equipment, set up the circuit seen above using the red & 3V sockets of the D-battery voltage source, and record your observations.

1. How does the current through the bulb in a single-bulb circuit compare with the current through the same bulb when it is connected in series with a second bulb? What does this imply about the current through the battery?

2. Compare the brightness of the two bulbs in the two-bulb circuit with each other. What can you conclude from this observation about the amount of current through each bulb?

3. It’s possible that some affects you observe might be due to manufacturing irregularities in the bulbs – that so-called “identical” bulbs might not, in fact, be quite identical. How might you test whether any differences are due to manufacturing irregularities?

D. On the basis of your observations and the reasoning you used above, respond to the following questions:

1. Is current “used up” on the first bulb, or is the amount of the flow the same through both bulbs?

2. Can you tell the direction of the flow through the circuit?
3. How does the amount of the current flow through the battery in a single-bulb circuit compare with the flow through the battery in a circuit with two bulbs connected in series?

E. Now consider the circuit pictured below with the circuit branching to two bulbs and back again. (Two bulbs connected in this way are said to be in parallel.) Predict with your partner(s) how the brightness of the two bulbs will compare to the brightness of an identical bulb in a single-bulb circuit.

F. Predict how the amount of current flowing through the wire at point 1 relates to the amount of current flowing through the wires at points 2 and 3?

G. Gather the needed equipment, set up the circuit seen above using the red & 3V sockets of the D-battery voltage source, and record your observations.

1. How does the current through the bulb in a single-bulb circuit compare with the current through the same bulb when it is connected in parallel with a second bulb?

   What does this imply about the current through the battery, i.e., the current through the wire at point 1?

2. Compare the current through the battery in the series circuit with the current through the battery in the parallel circuit. Is it lower, the same, or higher?

1. Thinking of the bulb in this way, would adding more bulbs in series cause the total obstacle to the flow, or total resistance, to increase, decrease or stay the same as before?

2. Formulate a rule for predicting how the current through the battery would change (i.e., whether it would increase, decrease, or remain the same) if the number of bulbs connected in series were increased or decreased.
3. Thinking of the bulb in the same way, would adding more bulbs in parallel cause the total obstacle to the flow, or total resistance, to increase, decrease or stay the same as before?

4. Formulate a rule for predicting how the current through the battery would change (i.e., whether it would increase, decrease, or remain the same) if the number of bulbs connected in parallel were increased or decreased.

References for worksheet

APPENDIX B: EXAMPLE DIALOGUE – GTA 13 ROUND 1, SESSION 2
The following is snippet of dialogue between GTA 13 and the virtual class during their simulator session. The context of the session included avatar-students working in groups to collect and analyze data to determine the period of a pendulum, a lab from the “Thinking Critically” curriculum. Avatar-students CJ and Ed worked in a group while Kevin, Maria and Sean worked together.

GTA 13: Start classroom. [pause] Uh, so hi everybody. Let's start the class.
Sean: Hi.
GTA 13: Uh, how are you, how, what's going on?
Sean: Pretty good! How are you?
GTA 13: Good. Thank you. So, I know that you guys completed the experiment. So how about tell me, uh, you guys, how did you do the experiment? Sean, you want to start?
Sean: Uh, sure. Uh, where do you want me to begin with like technique or results or where do you wanna start?
GTA 13: The procedure. How did you precede the experiment?
Sean: Okay. So, um, what we did was we have, of course we have to measure at two angles at 10 degrees or 20 degrees. So, what we did was we measured 20 swings five times at 10 degrees, and then we measured it 20 swings, five times at 20 degrees.
GTA 13: Alright. So, how about the other group? CJ, How did you measure the time?
CJ: Um, well we measure 10 at 10 degrees and 10 at 20.
GTA 13: All right. So, we have two different, uh, periods. So, CJ and Ed, measured time, 14, 10 code 10 swings and Kevin and Maria, Sean, they measured the time for 20 swings. Right? So.
Kevin: Yeah, we did a lot more than they did.
GTA 13: So, Oh yeah, yeah. If you really think so. Oh yeah, it’s double, right? Uh, so, uh, uh, Kevin. So, are you thinking, do you think that your data are reliable?
Kevin: Um, well we've got a pretty high t score, so I'm thinking our data. There must, there must be something wrong with it.
GTA 13: Okay. So, what is the t-score?
Kevin: The t-score we got was 3.71.
GTA 13: Okay. So, CJ and Ed, do you think your data are reliable, more reliable than them?
CJ: Yeah.
GTA 13: Uh, how do you determine your data is more reliable?
[The GTA continued to interact with avatar-students until the timer went off.]
Facilitator: Stop Classroom. ... You just got that one in [Stretch-It question]. Awesome! Why don't you give that [microphone] to <other GTA> and while you are doing that, how did you think that one [round] went?
GTA 13: Actually, I think I got some questions in this time. I asked about their assumptions. My students always forget about the assumptions...
[Discussion continued]
APPENDIX C: DIALOGUE EXAMPLE - GTA 6 USING EXPLAIN LOGIC
IN SESSION 3
The following is a snippet of dialogue between GTA 6 and the virtual class during their simulator session focused on group dynamics. The context of the session included avatar-students working in groups to predict and discuss lightbulbs in three different circuit orientations. In this excerpt, GTA 6 checked in with Sean, Kevin, and Maria about their progress and found the three avatar-students might have disagreed with each other, causing group dysfunction.

GTA 6: What, what part did you guys get to?

Sean: Um, well do you mean like as a group?

GTA 6: Yes. And not individually cause it’s group work.

Um, well, yeah, I mean, we made our predictions separately, but we got up to I think, let me look, let me check my work. Okay. Kevin and I got through, up to like right before and then, um, we, I mean, yeah, we kind of, I think we worked through most everything. I mean, Maria was kind of doing the right thing cause, uh, I don't know. She just seemed like she knew she was doing it. Um, she wasn't really great with any of our predictions. And so, she has kind of like went off on her own.

GTA 6: So, so you guys didn't work with her because she didn't agree with what you guys are predicting? Is that what you're saying?

Sean: Well, it's not like we didn't want to work with her. It's just that we were making predictions and then she was like saying like that they weren't right. So, then she, like, we were trying to discuss it with her, but she, I don't know. We just, I don't think we were, we didn't really communicate really well I guess.

GTA 6: I see. Well, the whole emphasis of the-

Sean: She just like set the whole thing up.

Well, well, the whole purpose of the group work is for you guys to work together and discuss things and you don't have to be right, or you don't have to be wrong, and you just have to, um, see what each other is thinking. And then with the predictions, there's no right or wrong. It's just something that you come up with based off of what you feel that should be or off of some type of evidence that you have.

GTA 6: Yeah. I mean, we were saying that, um, you know, to Maria, like we should talk about our predictions and stuff, but I don't know. I mean, I don't want to talk, I don't want to speak for Maria, but I just felt like she just, wasn't, didn't want to hear [pause] [p1802: Well, well...]

Kevin: what we had to say. She felt like we were wrong.

Well, Maria, let me ask you, why did you think that their predictions are wrong? Or why did you feel the need to say that their predictions were wrong?
APPENDIX D: BOTH COHORTS INDIVIDUAL GTA USE OF FOLLOW-UP IN THE CLASSROOM
Table 11: GTAs’ Use of Stretch-It: Follow-Up in the Real Classroom.

<table>
<thead>
<tr>
<th>GTA</th>
<th>Observation by Lesson</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unit 4</td>
</tr>
<tr>
<td>I</td>
<td></td>
</tr>
<tr>
<td>II</td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>-</td>
</tr>
<tr>
<td>IV</td>
<td></td>
</tr>
<tr>
<td>V</td>
<td></td>
</tr>
<tr>
<td>VI</td>
<td>0</td>
</tr>
<tr>
<td>VII</td>
<td>-</td>
</tr>
<tr>
<td>VIII</td>
<td>0</td>
</tr>
</tbody>
</table>

In the table above, shading indicates the observation included Follow-Up. A dash indicates the GTA did not have an observation that week. *Partial Data Collected. †Outlier observation.
Pre-Interview
A. Welcome Script: Hi, my name is [insert name here]. [Ask names of participants.] Thank you for your participation!
B. Introductory Narrative: As part of a larger research project to improve STEM instruction here at UCF, I am interested in hearing about your experience as a student in the introductory-level physics or chemistry lab. This interview could last up to one hour and will be audio-recorded. If you do not feel comfortable answering any questions or if you need to leave this interview early, you will not be penalized, and you will still be compensated $15. Do you feel comfortable with being audio-recorded? Do you have any questions before we get started with the [group] interview?
C. Informed Consent: Share Explanation of Research
[Give the form to the students and give them a minute or two to read the form. Then, briefly walk them through the form.]
Do I have your consent to audio-record this interview?
[Talk about norms here.]

**Topic Domain I: Student Reactions to Current GTA Interactions**

1) What does a typical day in your lab look like?
2) What interactions do you have with your TA during lab?
   a. Could you describe that in more detail?
3) What kinds of activities do you do in class?

**Topic Domain II: Student Experience with Group Work**

1) In your lab/class, do you work in groups or individually?
   a. How many people were in your group?
   b. Can you describe how each of your group members contributed to the group work?
   c. What was your role in your group?
   d. How do you, as a group, figure out who did what?
2) Do you think group work supports your learning? (If yes, in what ways? If not, why do you think it does not help?)
   a. What do you like about group work?
   b. What do you not like about group work?
3) Do you think group work helps you learn?
**Topic Domain IV: Group Work Prompts**

Below are three scenarios of students working in groups on an activity in either their classroom or lab. Please read through each scenario and think about if any of these scenarios feel familiar or relatable.

Scenario 1:

Miguel is working in a group with three other students on a lab activity. Miguel has assumed his role as group leader since his ideas have always been correct from the start of the semester. Everyone in the group seems to go along with his idea and each contributes to the discussion about the experiment except for one student, Sophia. Sophia remains quiet during the discussion. When it is time to conduct the experiment, she writes down everything the group says in their lab report or lab notebook. She also does not touch any of the lab equipment since Miguel is always the first one to start setting up the experiment. After the lab is finished, Sophia looks frustrated.

Scenario 2:

During recitation, Ray is working with two other students on a worksheet about energy. After working through the first section, Ray finds themself confused about something on the second page. Ray has also noticed that it has been about twenty minutes and no one in their group has said anything to each other. Ray looks over and sees one student is far ahead in the worksheet and the other student is on their phone.

Scenario 3:

Liam, Emily, and Charlie all have different ideas about the relationship between force and motion. Liam thinks his idea is correct and will not listen to what Emily and Charlie have to say. Emily and Charlie split off into their own group of two while Liam works alone. However, Emily and Charlie eventually become confused. Seeing that Liam has worked through the entire activity, they decide to go along with Liam’s ideas because they do not want to get a bad grade.

1. Did you have a similar experience as in the scenarios described above?
2. In what ways?
   a. For physics students: Did this happen in the lab or recitation or both?
3. How did you feel?
   a. How did this situation impact your feelings of anxiousness in your classroom?
4. Were you comfortable sharing your ideas about the concepts or experiment with your peers in this situation?
   a. Why or why not?
5. Did your TA notice or acknowledge the situation?
6. What did your TA do or say? (or How did your TA react or respond?)
   a. How did that make you feel?
b. Did your TA’s actions help reduce or increase your feelings of anxiousness?

7. Do you think your TA’s actions helped?
   a. Was the issue resolved? Why or why not?

8. What do you think a TA could do to be helpful with resolving an issue like this?

9. Were there any other challenges your group encountered that are not described in the scenarios?
   a. Did your TA notice the challenges you were encountering?
   b. What did your TA do in this situation?
   c. Can you give an example of what your TA said or did to help your group work well?
   d. Do you think what your TA did was helpful in that situation?

10. Do you have any suggestions about how a TA can help students in groups share their ideas with each other? (or help with a disagreement?)

**Topic Domain V: Student Perception of Examples of Facework**

We have a few examples of how a TA might help students feel comfortable with sharing their ideas in front of their peers (give list to students to read):

1) Works to avoid making you look bad
2) Gives positive verbal feedback
3) Shows that he/she cares about your learning experience
4) Shows respectful interest in your ideas
5) Seems attentive to you as an individual

1) Which one(s) have your TA used in your class/lab? Can you give an example of how they did that?
   a. Are there any strategies you wish your TA used to help you feel comfortable sharing your ideas in front of your peers?

**Topic Domain VIII: Student Perception of GTA’s Questioning Skills**

1) How frequently does your TA ask questions in your class?
   a. Do they do this in front of the whole class, in front of small groups, or both?

2) What do you think is the purpose of your TA asking you questions in class?

3) Do you think the questions help you learn?
   a. Do you feel like you are being interrogated?

4) Tell me more about why you think the questions help you learn or why you don't think they help.
5) Do the questions your TA asks you help your thinking or reasoning during lab or recitation?

**Topic Domain VI: Student Experience with Cold Calling and Error Framing**

1) Does your TA call on non-volunteering students or groups to answer a question?
   a. How frequently does your TA do this?
   b. Do they do this in front of the whole class, in front of small groups, or both?
   c. Do you like being called on?

2) How do you feel when you are called on to share your ideas in front of your peers?

3) Do you feel like your idea was valued by your TA and your peers?

4) How do you think being called on impacts your learning?

5) How did your TA respond when you shared an answer that was not completely correct, or when your answer was not what your TA was looking for?
   a. Could you give an example of what they would say?

6) How did those responses make you feel?
   a. How did those responses help you value reasoning over getting the correct answer?

**Topic Domain IX: Student Perception of the Impact on Feelings of Anxiousness**

1) Are there any activities in your lab that may increase or decrease your feelings of anxiousness?
   a. Why or why not?

2) How does working in groups on an activity in your classroom/lab impact your feelings of anxiousness?

**Topic Domain VII: Promoting a Welcoming Environment and Student Perception of Decreasing Negative Evaluation**

1) Can you describe what your TA has done to make you comfortable with sharing your ideas with the whole class or your lab group?
   a. How can a TA respond to an incorrect student answer to make you feel like your ideas are valued?

2) Do you have any suggestions on how a TA can help their students feel more comfortable sharing their answers and ideas with their peers?

**Topic Domain X: Student Perception of GTA’s Role**
We sometimes describe classes as either focused on "getting the right answer" or "understanding how course concepts work together and can be applied to new scenarios". Which do you think is more descriptive of your class?

There are a lot of different ways a TA could teach a class/lab. Examples include talking at the board, helping with group work, asking questions, etc.

1) What does your TA typically do during lab/recitation?
   a. Why?
   b. How would you describe your GTAs actions in the lab?
   c. Do you consider them an instructor or a mentor?
APPENDIX F: STUDENT QUOTES DESCRIBING GTA ROLE – MENTOR OR INSTRUCTOR
Student descriptions of Mentor vs Instructor

**Mentor descriptions:** relatable, friendly, comfortable to interact with (ask for help), helped on a personal level, checked in with students, guiding, took time with student interactions

Student C1: “Um definitely more like a mentor. I didn’t. I think we all felt very comfortable with him.”

Student D1: “…he was like checking on us and stuff like that, and monitoring our progress, especially doing like, lab and stuff that part more so goes into like the mentorship thing.”

Student F1: “It's like, when you go to like a professor's class, you feel very like, okay, you know, like, I gotta behave in a certain way and do whatever and, you know, be like, very respectful, not that I wouldn't be respectful to her, but just in a different way, it was more of like a friend almost. And someone that, you know, I felt comfortable coming to, to ask for help, or, yeah, I felt more of like, she was like a mentor.”

Student F2: “I feel like she wasn't really just talking at us all the time. She really would take her time to speak to us individually and see how she, she can get the point across to each of us individually, as opposed to just like, just telling us all the facts.”

Student G1: “…as far as like, on a personal level, like helping us, I would say as a mentor.”

Student H1: “To me, I see mentors as you know, somebody who’s looking over your shoulder and guiding you along the way and helping you out whereas an instructor is that person at the front of the class, you know, lecturing you to death. And it’s just very unenjoyable.”

Student H2: “She’s just relatable. I don’t know. I she just I don’t know, she makes you feel comfortable, at least for me. She was comfortable. She’s like, it’s alright. I understand. She’s like when I was learning this stuff. She’s like, it wasn’t easy for me. You know, she always says, you know, guys, I studied a little more definitely when I was learning was, um, was confusing.”

**Instructor descriptions:** talked about lab content only, punctual, explained, helped in professional manner, did standard procedures, teacher

Student A: “I consider him more of an instructor. Because he didn't really talk about physics outside of like, how it would relate outside of the lab, or outside of the recitation.”

Student B2: [student wished their TA was mentor-like] “I mean, as far as like making sure that we were very punctual with our time as far as like going from recitation, like what the worksheet to the quiz to then the lab, she definitely did pretty well. But I feel like to make her more of a mentor is to make her a little bit more relatable. You know, we already get a lot of
instruction from her, regular professors. So, you know, I feel like going to lab should be more comforting and a little bit more relatable, especially because the TA is a little bit more closer in age and pretty much went through the same thing that we went through already. So, I feel like that would have definitely made her more of a mentor if she did that.

Student D1: “...there’s parts where he was just like, explaining things like a professor…”

Student E1: “I don't think we were not, I mean, at least, in my experience, we weren’t on a personal level like that. He really helped me to understand concepts in a professional manner. Not that a mentor isn’t professional, but we never like discussed our life or, like, what I'm doing in school or anything like that we solely talked only about physics in that class. And he also did a good job at maintaining, like, good, professional professionalism, as the TA of the class, but yeah, definitely an instructor.”

Student G1: “Yeah, I'd say more instructor in the sense that, like I said, he taught the lab and did the standard procedures as far as like the recitation quizzes and everything…”

Student H2: “But she is still like a GTA. So, um, I do see her as an instructor because like, she’s our teacher for that time.”

Other – Instructor and Facilitator – Facilitator description: helpful

Student C2: [Described as a facilitator as well] “Um, I think a little bit of a[n] instructor and like, a facilitator as well, like just helping us. Like, know, how to set up the material and like, kind of help us understand what we’re doing or help us understand, like, what we how should we should be thinking.”

Other – Student Teacher descriptions: student, TA as a secondary role, did not care about students

Student E2: “Like not a mentor, not even an instructor, like he's barely making that mark, in my opinion. I would say like, the way this GTA was perceived to me, because like, I know that he's a graduate student, so he's already going to be taking classes like all of us, that it was, like, the way I felt, was, this was something added to everything else that he was doing, that he didn't want to do. And clearly, he showed that he didn't care. And so, he's just the student teacher”

Student G2 and I1 did not give a description of how or why their GTA was a mentor or instructor. Student B1 did not answer the question about their GTA’s role.
APPENDIX G: IRB APPROVAL
EXEMPTION DETERMINATION

February 10, 2021

Dear Constance Doty:

On 2/10/2021, the IRB determined the following submission to be human subjects research that is exempt from regulation:

<table>
<thead>
<tr>
<th>Type of Review:</th>
<th>Initial Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title:</td>
<td>Optimizing Mixed-reality Simulation to Support Physics Graduate Teaching Assistants’ Development of Student-centered Pedagogical Skills</td>
</tr>
<tr>
<td>Investigator:</td>
<td>Constance Doty</td>
</tr>
<tr>
<td>IRB ID:</td>
<td>STUDY00002689</td>
</tr>
<tr>
<td>Funding:</td>
<td>None</td>
</tr>
<tr>
<td>Grant ID:</td>
<td>None</td>
</tr>
</tbody>
</table>

This determination applies only to the activities described in the IRB submission and does not apply should any changes be made. If changes are made, and there are questions about whether these changes affect the exempt status of the human research, please submit a modification request to the IRB. Guidance on submitting Modifications and Administrative Check-in are detailed in the Investigator Manual (HRP-103), which can be found by navigating to the IRB Library within the IRB system. When you have completed your research, please submit a Study Closure request so that IRB records will be accurate.

If you have any questions, please contact the UCF IRB at 407-823-2901 or irb@ucf.edu. Please include your project title and IRB number in all correspondence with this office.

Sincerely,

Katie Kilgore
Designated Reviewer
LIST OF REFERENCES


This is a virtual network that provides resources for biology GTAs and trainers (www.biotap.org).


J. A. Luft, J. P. Kurdziel, G. H. Roehrig, and J. 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).


[37] A. E. Austin, Preparing the next generation of faculty: Graduate school as socialization to the academic career. The journal of higher education, **73**(1), 94-122 (2002).


in the cross-institutional evaluation of graduate teaching assistant professional development programs. CBE—Life Sciences Education, 17(1), ar8 (2018).


https://www.aip.org/sites/default/files/statistics/graduate/1styeargradrpt-1516.pdf


[77] Y. Cao, C. Smith, B. D. Lutz, and M. Koretsky, Cultivating the next generation: Outcomes from a Learning Assistant program in engineering. in


[108] T. A. Roman and S. Uttamchandani, Researching pedagogy within small active learning classrooms: Examining enacted pedagogies of learner and


[112] Mini-studio sections are paired with lecture such that a portion of the lecture grade is attributed to students' performance in their mini-studio section.

[113] Students who take these courses are non-majors, with majority of the students planning on attending medical school after they graduate.


observation protocol for undergraduate STEM. Journal of Chemical Education, 93(7), 1191-1203 (2016).


[121] We observed two GTAs in Fall 2018 teaching the same lesson twice, and thus have taken the average of the two observations as one observation.


[123] Four Physics I GTAs practiced Stretch-It during Round 1 and all six Physics I GTAs practiced Stretch-It in Round 2. All five Physics II GTAs practiced Stretch-It during both rounds.


[126] It is possible conducting student interviews during the pandemic limited the students who participated in this study since students with illness or working in health care might not have been able to participate.


