INVESTIGATING NEWER STATISTICS INSTRUCTORS' BREAKTHROUGHS WITH AND MOTIVATIONS FOR USING ACTIVE LEARNING: A LONGITUDINAL CASE-STUDY AND A MULTI-PHASE APPROACH TOWARDS NSTRUMENT DEVELOPMENT

by

Elijah Sterling Meyer

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

Doctor of Philosophy

in

Statistics

MONTANA STATE UNIVERSITY
Bozeman, Montana

July 2022

DEDICATION

I would like to dedicate this dissertation to my mother, Rita. Your unwavering love and support shown to me through the difficult times always gives me the strength to continue pushing forward. At an early age, you taught me the importance of teaching and education. It is your guidance that has helped inspire me to become a teacher, like you. You continue to be an inspiration to not only myself, but others that cross your path. I love you the mostest.

ACKNOWLEDGEMENTS

I would like to sincerely thank my advisor, Jennifer Green, for her support, guidance, and willingness to listen. I feel extremely lucky to work with someone that continuously brings out the best in me and my work. I cannot thank you enough for the countless hours invested in my growth as a researcher, scholar, and individual. I additionally want to extend a thank you to my co-advisor Stacey Hancock, Andy Hoegh, Mark Greenwood, and Megan Wickstrom. The opportunities each of you has provided me throughout my journey have helped shape who I am today as a researcher and educator. I cannot thank each of you enough for your overwhelming amounts of guidance, support, and encouragement. Your willingness to chat about research and teaching is extremely appreciated and is how I strive to work with students in the future.

I want to thank Christian Stratton for being the friend that I needed during graduate school. The combination of your kindness and support created for some unbelievable memories. You are one cool cat. I want to additionally thank Justin Gomez. The overwhelmingly sarcastic, yet sincere, support from a true friend like yourself is something I have and will continue to cherish. I feel lucky to have made so many memories with you along the way and can't wait to continue doing so. Finally, I would like to thank Mandy LeBrun for her support and continuously assuring me that everything was always going to be okay.

I additionally want to thank Jade Schmidt and Melinda Yager for their investment and support towards my development as a teacher. Without the support of these people and others, I would not be where I am today. I am truly grateful for my experiences with you all and cherish the relationships we have built.

TABLE OF CONTENTS

TABLE OF CONTENTS - CONTINUED

TABLE OF CONTENTS - CONTINUED

TABLE OF CONTENTS - CONTINUED

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES - CONTINUED

ABSTRACT

National recommendations call for a shift from using lecture-based approaches to using approaches that engage students in the learning process, primarily through active learning techniques. Despite these recommendations, the adoption of active learning techniques for newer statistics instructors remains limited. The goal of this research is to provide a more holistic understanding about statistics instruction, specifically as it relates to recommended active learning techniques and newer statistics instructors, including graduate student instructors (GSIs).

In this research, I present two studies. In the first study, we investigated GSIs' breakthroughs in their knowledge about, emotions towards, and use of active learning over time by using a longitudinal collective case-study approach. Survey, interview, and observation data across four semesters revealed that the GSIs' breakthroughs in their use of active learning only occurred after their increased knowledge about active learning aligned with their emotions towards it. This study further revealed that the GSIs needed to feel confident in and be challenged by their course structure before implementing active learning techniques.

The second study builds upon these findings by exploring statistics instructors' motivations or reasons for using active learning. Under the self-determination theory framework, we conducted a multi-phase study to develop an instrument that measures four different types of motivational constructs for using group work, a specific active learning approach. We constructed items using expert opinion and cognitive interviews, and then we conducted two pilot studies with newer statistics instructors. The resulting reliability and validity evidence suggest that this instrument may help support future studies' investigations of motivation, helping us to better understand newer statistics instructors' use of active learning. Together, these studies may help inform future recommendations on how to support newer statistics instructors' early adoption of such techniques.

INTRODUCTION

Statistics instruction is changing. Recently, national guidelines have recommended the use of more student-centered teaching techniques, such as active learning, in statistics classrooms to better engage students in the learning process (Carver et al., 2016; Conference Board of Mathematical Sciences, 2016). When learning through such techniques, students have shown increased examination rates, increased retention of material, and lower failure rates versus those learning through more lecture-based instruction (Freeman et al., 2014). Despite this, the call for active learning has been met with mixed responses, especially from newer statistics instructors, such as graduate student instructors (GSIs) who serve important roles in teaching undergraduate statistics courses (Auerbach et al., 2018; Justice, 2020). This gap between recommended and enacted instructional practices presents a need to better understand newer statistics instructors' experiences with and use of active learning. This dissertation presents research that addresses this need by investigating how GSIs breakthrough with active learning, and why newer statistics instructors, such as GSIs, use active learning when teaching statistics through the lens of motivation. Our research aims to help inform future recommendations on how to support newer statistics instructors' early and consistent use of active learning techniques.

Statistics Instruction

Data are everywhere. The information citizens see daily has more than doubled since 2010 and is projected to continue increasing by roughly 5% each additional year (Heim & Keil, 2017). This large influx of data into everyday life has emphasized the need for a statistically literate society. Because of this need, it is widely recognized that statistics is among the most important quantitative subjects available through university curricula, being required in over 100 different college majors (American Statistical Assocciation, n.d.; Weihs & Ickstadt, 2018). In response, statistics instruction is advancing to ensure individuals within these courses are prepared to make meaning of the data around them.

The Guidelines for Assessment and Instruction in Statistics Education (GAISE) College Report (Carver et al., 2016) recommends that statistics instructors emphasize statistical thinking, foster active learning and conceptual understanding, and integrate the use of real data and technology to explore data. Recently, there has been a concerted effort to develop techniques that answer this call to improve and advance the teaching of statistics. These techniques include the creation of applet interfaces to create simulations (e.g., Rossman & Chance, 2021), programs that simulate real world data collection (e.g., AMSPP, 2021), and reform efforts to emphasize statistical literacy and develop statistical thinking (e.g., Wood et al., 2018). Common to all these developments is their facilitation of active learning in the classroom (Wood et al., 2018).

Active Learning

Active learning is defined as any instructional method that engages students in the learning process (Bonwell & Eison, 1991). The core element of active learning is student engagement, which is often contrasted to the traditional lecture where students passively receive information from the instructor (Prince, 2013). There is a plethora of research to suggest that learning through the use of active learning techniques is valuable for students. A meta-analysis of 225 studies revealed evidence that STEM students learning through active learning techniques earned higher exam scores and had lower failure rates than those learning under more traditional lecture styles of teaching. On average, students' exam scores were reported to be about 6% higher and students were 33% less likely to earn a failing grade in courses that used active learning than in courses that didn't (Freeman et al., 2014). Other studies have also found meaningful increases in student exam scores (Cotner et al., 2013), with average increases up to 12% with the use of active learning (Gross et al., 2015). Active learning is further connected to improved communication, community, connectedness, satisfaction, and flexibility (Allsop, Young, Nelson, Piatt, & Knapp, 2020). As research continues to uncover how the use of active learning can benefit students, it's crucial that statistics instructors align their practices with current recommendations to use active learning when teaching statistics.

Statistics Instructors' Challenges with Active Learning.

Despite recommendations to use active learning in the classroom and the growing recognition of active learning's benefits, most college STEM instructors still choose traditional, instructor-centered teaching methods (Deslauriers, McCarty, Miller, Callaghan, & Kestin, 2019). When asked, instructors identified many challenges that prevent them from adopting active teaching strategies, including insufficient time, limited resources, a lack of departmental support, concerns about content coverage, and concerns about evaluations of their teaching (Deslauriers et al.). Many instructors also shared perceptions that students resist active teaching strategies and prefer traditional methods. As a result, these instructors often reverted to using more passive modes of instruction such as lectures (Deslauriers et al.).

The use of active learning may be even more challenging for newer instructors, such as a graduate student instructor (GSI). We define a GSI as any graduate student who is the lead instructor of a statistics course or statistics recitation section. With often little to no teaching experience, statistics GSIs frequently hold beliefs that are not aligned with current recommendations for teaching statistics and have not reviewed materials on how to teach statistics (Justice, 2020). This makes it challenging for them to use potentially new and unfamiliar active learning techniques to enhance learning opportunities. Justice et al. (2017) found that when statistics GSIs are first introduced to active learning, there is no consensus among them about how to use such strategies to deliver course content.

In addition to the challenges faced as newer statistics instructors, GSIs may have little incentive to use active learning. A national sample of graduate students reported that over 70% of graduate students encountered a stressor that interfered with their well-being (El-Ghoroury, Galper, Sawaqdeh, & Bufka, 2012). Studies have shown that poor stress management has a negative relationship with motivation in students (Park et al., 2012). The additional workload and time investment to learn how to use more advanced teaching techniques may be stressful and decrease GSIs' motivation to do so.

Despite these challenges, GSIs have a prominent role in statistics education. Many statistics departments rely on graduate students to teach courses (Justice, 2020), and in the United States alone, GSIs teach almost 50% of introductory statistics courses for a variety of students at four-year colleges and universities (Blair, Kirkman, Maxwell, & Society, 2013). With the emphasis on active learning in many introductory statistics curricula, it's important to understand statistics GSIs' experiences with active learning and how they can be supported in their use of active learning. Currently, research on newer statistics instructors focuses largely on the challenges GSIs face when teaching and recommendations to best prepare GSIs to teach (e.g., Garfield & Everson, 2009; Justice 2020; Mongillo, 2016). However, there is still much to learn about newer statistics instructors' relationship with active learning.

Research Purpose

The purpose of this research is to provide a more holistic understanding about newer statistics instructors' relationship with active learning. Existing research has positioned the use of active learning techniques in statistics education. The findings of Freeman et al. (2014) highlight the academic benefits students receive when learning through such techniques, making it crucial that all types of statistics instructors, experienced and newer, adopt active learning techniques into their statistics classes. Yet, there still exists a gap between recommendations on how to teach statistics using active learning and the practice of teaching statistics (Tishkovskaya & Lancaster, 2012). Identifying challenges from existing literature, this research provides unique insight into the evolutionary process of GSIs' knowledge about, emotions towards, and use of active learning. We believe that investigation around challenges related to GSIs' knowledge about and emotions towards active learning, and how they overcome these challenges could inform future recommendations on how to support newer statistics instructors' early adoption of such techniques.

Additionally, this research poses the argument that newer statistics instructors may be more willing to engage and persist with the use of active learning if it is considered a self-determined behavior, i.e., if they make their own choices and decisions to engage in the use of active learning. Motivations, the reasons why we engage in behavior, are windows to explore how self-determined a specific behavior may be. Encouraging newer instructors towards the use of active learning starts with a detailed understanding of what motivates them as instructors to use active learning. For example, an instructor can be motivated to use active learning in hopes of receiving a teaching award or be motivated to avoid undesirable outcomes such as poor student reviews. These types of motivations would suggest that the use of active learning is not very self-determined and may result in shorter, less desirable uses of active learning. In contrast, an instructor may use active learning because they value and enjoy active learning as a method of teaching. This type of motivation characterizes more self-determined behavior and is associated with long-lasting, self-sustaining engagement. In general, different types of motivations are associated with varying degrees of persistence towards a behavior (Ryan & Deci, 2000); departments should encourage newer statistics instructors to cultivate certain types of motivation that are associated with greater persistence in the use of active learning.

Motivation can be investigated though the use of instruments. Although there exist instruments to measure motivation, there are no known instruments that measure newer statistics instructors' motivations to use active learning. Given this need, our research sets out to initiate the development of an instrument that measures newer statistics instructors' motivations to use active learning. This instrument may help support future studies' investigations of motivation, helping us to better understand newer statistics instructors' use of active learning. A better understanding of newer statistics instructors' use of active learning can inform how to support the use of active learning as a more self-determined behavior.

Research Journey

Positionality

My experiences at Montana State University motivate my work in exploring newer statistics instructors' relationships and experiences with active learning. As a statistics GSI at Montana State University, I experienced many of the challenges discussed above when trying to first use active learning in my classroom. As I continued teaching, I discovered my appreciation for such techniques. I especially enjoyed the learning communities I was able to foster among students through the use of active learning. As I saw the benefits of active learning play out in my classroom, my research focus started to take shape around the use of such techniques.

After gaining teaching experience, I was involved in the professional development program for GSIs in the Department of Mathematical Sciences at Montana State University. One aim of this program was to best prepare incoming graduate students for their first day of teaching. I gained experience in mentoring, facilitating workshops, and creating activities to help support graduate students in teaching, including in their use of active learning. The conversations I had with both GSIs and colleagues during this time further shaped my research trajectory, and I knew I wanted to continue working with newer statistics instructors, such as GSIs, and help support their use of active learning.

As my research trajectory took shape, I developed beliefs that GSIs are the future of introductory statistics education, and that all newer statistics instructors, including GSIs, can incorporate active learning techniques into their classrooms with the proper support. These beliefs motivated the following studies presented in this dissertation. Through this research, I hope to expand our current understandings about newer statistics instructors' relationships and experiences with active learning.

Chapter 2: Understanding How and When Graduate Student Instructors Break Through Challenges with Active Learning

Positionality In Chapter 2

My experiences as a GSI mirror those of my two participants in Chapter 2. At the start of my master's degree, I attended and completed the same professional development training program. After this, I was immersed in similar weekly teaching meetings that our department offered and attended similar weekly course-coordinated meetings. During the recruitment phase of this research, I was involved with the planning, preparation, and implementation of teaching workshops both participants attended. Because of this, I built a trusting relationship with each participant. While building this trust, I was able to understand their backgrounds more clearly, which informed the selection of each individual for this study. The relationship I developed with each participant could have influenced the data collection process, resulting in response bias towards more positive answers during interviews about active learning. However, the importance of this research was made clear, and it was fully understood that their answers were to be blinded, not reported to course supervisors, and never personally held against them by me or another researcher in hope to promote the most honest answers possible. Any qualitative data analysis in Chapter 2 was a collaborative effort to ensure that my beliefs about active learning, my beliefs about statistics GSIs, and my relationships with each participant did not influence the results.

Description

With an interest in newer statistics instructors' use of active learning, we set out to understand how GSIs break through challenges with active learning. Through existing literature, we identified challenges for newer statistics instructors, such as GSIs, to use active learning. These challenges were related to their knowledge about and emotions towards such techniques.

Because we defined a breakthrough as a positive shift in emergent themes across time points of data collection, the design of this study was longitudinal to investigate when breakthroughs occur. We used purposeful sampling to select one GSI that was classified as a lecturer and one that was classified as a facilitator. We used a collective case-study to understand how each GSI experienced breakthroughs with active learning. The data collection process was framed in a post-positivism framework, where data collection procedures were planned, consistent, and similar across our two participants at each time point. Across four semesters, we collected three different types of data: survey data, interview data, and video data. With participants' trust earned prior to the start of the study, the epistemological assumption was that each participant would share answers to interview questions truthfully and fill out survey questions honestly.

We identified breakthroughs when data collected at different time points signified evidence towards a positive shift in knowledge, emotions, or use of active learning. Breakthroughs were found through a constructivist framework. We did not begin with prior theories regarding either GSI's knowledge about, emotions towards, or use of active learning to help explain results, but rather generated patterns of meanings through the story these data told (Cresswell, 2003). Interview and survey data were analyzed using a provisional coding approach. We created a video observation protocol to guide how we understood each GSI's use of active learning. Survey, interview, and observation data revealed the importance of aligning GSI's increased knowledge about and emotions towards active learning before they could experience a breakthrough in their use of active learning. This study further revealed that GSIs need to feel confident in and be challenged by their course structure before implementing active learning techniques. The following chapters build upon these findings by proposing a framework and developing an instrument to measure the reasons why newer statistics instructors use active learning.

Chapter 3: A Framework for Understanding Statistics Instructors' Motivations to Use Active Learning

In Chapter 3, we discuss and justify the use of self-determination theory (SDT) (Ryan & Deci, 2000) as a motivational framework to investigate the reasons why statistics instructors use active learning. SDT theorizes that more self-determined behavior is associated with more positive behavioral outcomes. An advantage of this framework is its emphasis on how statistics instructors' use of active learning can become more self-determined. For an instructor's use of active learning to become more self-determined, SDT suggests that their needs for competence, relatedness, and autonomy must be fulfilled. We additionally incorporate self-regulation learning theory as a means to promote active learning as a more self-determined behavior.

Classifications of self-determined behavior are understood through three different types of motivation (intrinsic, extrinsic, and amotivation) and six subsequent regulatory styles (integrated regulation, identified regulation, introjected regulation, external regulation, non-regulation). We argue that studying statistics instructors' use of active learning through these motivations and regulatory styles is key to better understanding how to foster statistics instructors' early use of active learning. We further justify the investigation of statistics instructors' use of active learning through this framework by highlighting its strengths compared to another popular motivational model, expectancy-value theory.

Through this chapter, we describe the lens in which we view statistics instructors' motivations to use active learning and use this lens as our framework in Chapter 4 to initiate the development of an instrument that measures newer statistics instructors' motivations to use active learning.

Chapter 4: Motivation Matters: Developing an Instrument to Measure Newer Statistics Instructors' Situational Motivation to Use Active Learning in the Classroom When Teaching Statistics

Positionality in Chapter 4

Assumptions made about motivation and newer statistics instructors, as well as my experiences as a statistician may have had a direct influence on what we intend to measure and how we measure it. We assume that motivation to use active learning is quantitatively measurable and that all types of regulatory styles we intend to measure can be measured on newer statistics instructors who use such techniques. We also assume that the creation of an instrument that measures types of motivations will allow others to make statements about newer statistics instructors' motivations to use active learning.

My experiences as a newer statistics instructor, as well as my experiences working with them in a professional development setting, revealed many internal and external factors that consumed their time. This informed me to make time sensitive decisions when creating this instrument to better ensure the authenticity of these data collected. To do so, I decided to only measure a subset of the regulatory styles outlined in SDT to help improve the practicality of our instrument.

The relationships I've made through the statistics and statistics education communities may have influenced the recruitment phases for this research. For example, when conducting cognitive interviews, I knew and had an existing relationship with five of the six participants. These relationships could have influenced the data collection process, resulting in responses that may have not fully articulated issues with the items' interpretability. However, the importance of this research was made clear, and it was fully understood that honest answers were needed to ensure a better understanding of how newer statistics instructors may interpret each item.

Description

In Chapter 4, we developed a multi-phase study, influenced by Justice et al. (2017), to start the development of an instrument that measures newer statistics instructors' intrinsic motivation, integrated regulation, external regulation, and amotivation when using active learning. We define a newer statistics instructor as an instructor with three or less years of post-graduate statistics teaching experience. We chose these four motivational constructs based on their stark differences with how they characterize behavior as more or less self-determined. The goal of this multi-phase study was to ensure that we were measuring these four motivational constructs as intended and consistently by gathering initial evidence of construct validity, content validity, and reliability for our instrument in development.

Using the GAISE College Report (Carver et al., 2016) and expert opinion, we chose to write our items in the context of group work. In Phase 1, we drafted initial items based on our review of existing instruments in Psychology and Statistics Education. In Phase 2, we conducted two rounds of expert item review by sending out a survey to solicit expert opinion on how they interpreted the created items. Each expert was asked to read each item, describe how they interpreted the item, and provide suggestions to improve item clarity. Phase 3 was comprised of cognitive interviews to assess item interpretability in real-time. During this phase, six newer statistics instructors answered items while in a video call as I took notes and facilitated discussion around each item. In Phases 4 and 5, we conducted two separate pilot studies. We investigated irregular response patterns and reliability estimates, and we explored results from a confirmatory factor analysis to gather evidence of validity and reliability. The resulting validity and reliability evidence suggest that this instrument may help support future studies' investigations of motivation.

Conclusion

Finally, Chapter 5 concludes this investigation of newer statistics instructors' relationships and experiences with active learning and presents directions for future research, outlining how future studies can build upon our findings. We hope that results from these and future studies may promote discussion on how professional development programs can support newer statistics instructors' early adoption and effective usage of active learning techniques.

UNDERSTANDING HOW AND WHEN GRADUATE STUDENT INSTRUCTORS BREAK THROUGH CHALLENGES WITH ACTIVE LEARNING

Contribution of Authors and Co-Authors

Author: Elijah S. Meyer

Contributions: Designed study, collected data, performed qualitative analyses, interpreted results, and wrote manuscript.

Co-Author: Jennifer L. Green

Contributions: Designed study, collected data, discussed results and conclusions, and edited drafts of manuscript.

Co-Author: Elizabeth G. Arnold

Contributions: Collected data, discussed results, and edited drafts of manuscript.

Co-Author: Megan Wickstrom

Contributions: Discussed theoretical framework and edited drafts of manuscript.

Manuscript Information

Elijah S. Meyer & Jennifer L. Green & Elizabeth G. Arnold & Megan Wickstrom

Journal of Mathematics Teacher Education

Status of Manuscript:
X___ Prepared for submission to a peer-reviewed journal
____ Officially submitted to a peer-reviewed journal
____ Accepted by a peer-reviewed journal
____ Published in a peer-reviewed journal

Publisher: Springer

Abstract

Across recommendations for teaching undergraduate mathematics and statistics, instructors, including graduate student instructors (GSIs), are encouraged to implement techniques that actively engage students in the material. Research suggests that instructors' knowledge about and emotions towards using active learning can promote or inhibit their use of active learning in the classroom. However, little is known about how GSIs' knowledge about and emotions towards using active learning evolve over time. We present findings from a longitudinal case study following two GSIs within a department of mathematical sciences across four semesters. We discuss observed breakthroughs regarding their knowledge of, emotions towards, and use of active learning techniques. Findings highlight when and how GSIs evolve with active learning, promoting discussion on how to support GSIs' early use of active learning in their classroom. We provide suggestions for professional development programs and discuss future research practice when investigating GSIs' longitudinal development as instructors.

Introduction

In recent years, the distribution of university instructors has shifted to include more graduate student instructors (GSIs), many of whom are teaching for the first time. The American Association of University Professors (AAUP, 2018). Synthesized data collected in 2016 from the Integrated Postsecondary Education Data System, revealing that among all instructors at Research I universities, GSIs make up the largest percentage. In Ph.D. statistics departments, it's reported that GSIs teach roughly 25% of introductory courses in the United States (Blair et al., 2013). Ellis (2014) discussed that universities' instructional needs are being met in-part by an increasing number of GSIs. Thus, GSIs have a prominent role in teaching undergraduate mathematics and statistics courses (Speer et al., 2005).

Along with the shift in who is teaching undergraduate-level mathematics and statistics courses, there has been a shift in the recommended instructional practices for these courses. Over the years, national reports have encouraged the use of active learning in the classroom to maximize student learning. Active learning is defined as the process in which students engage in meaningful learning activities that promote conceptual understanding and critical thinking (Bonwell & Eison, 1991). Within the mathematical sciences, the Conference Board of Mathematical Sciences (CBMS) and the American Statistical Association (ASA) promote the use of active learning, stating that a "wealth of research has provided clear evidence that active learning results in better student performance and retention than more traditional, passive forms of instruction alone" (CBMS, 2016, p. 1). Despite these recommendations, many instructors, including GSIs, adhere to traditional teaching approaches that they themselves experienced as students, where students passively receive information from instructors (Deshler, Hauk, & Speer, 2015; Harmin & Melanie, 2006). Further research suggests novice instructors (like GSIs) may not be ready to use active learning when they teach for the first time because they are likely in survival mode, trying to make it through the semester any way they can (Beisiegel, 2019). Other research suggests that, even when novice instructors attempt to use active learning, it is implemented differently than intended, mitigating potential learning benefits to students

(Auerbach et al., 2018). To help facilitate early and consistent uses of active learning in mathematics and statistics classrooms, we need to better support GSIs in teaching.

The use of active learning in classrooms is complex and there exist many challenges that impede an instructor's effective use of active learning in the classroom. Specifically, research has shown that these challenges relate to instructors' knowledge about and emotions towards active learning. For example, instructors with less knowledge about, or more negative emotions towards using active learning may be more hesitant to use, and less effective when using active learning in the classroom (Finelli et al., 2014; Henderson et al., 2012). Better support may help GSIs break through challenges related to their knowledge about and emotions towards active learning to start using active learning earlier and more effectively. However, little research has explored when and how these breakthroughs occur. In this study, we focus on breakthroughs in the context of a GSI's knowledge about, emotions towards, and use of active learning. Our research is designed longitudinally to allow for natural investigation into how and when GSIs break through challenges related to active learning. We follow two mathematics GSIs across four consecutive semesters and investigate the following research question: *When and how do breakthroughs occur for GSIs in their knowledge about, emotions towards, and uses of active learning practices across multiple semesters of instruction?*

By investigating GSIs' breakthroughs with, and the relationships between their knowledge about, emotions towards, and uses of active learning, our research will help promote more encompassing information on how to best support GSIs in their use of active learning while teaching. This research can lead to more informed and differentiated training for GSIs, helping to promote, enhance, and accelerate their use of active learning in classroom instruction and thereby enhance their students' learning.

Literature Review

Although the definition of active learning is fluid and looks different for each instructor, it can be generally defined as the process in which students engage in meaningful learning activities and conceptually understand the material (Bonwell & Eison, 1991). Research has shown that the adoption, integration, and implementation of active learning teaching techniques can help promote student learning, achievement, and confidence in mathematics (e.g., Freeman et al., 2004). The CBMS (2016) calls on "institutions of higher education, mathematics departments and the mathematics faculty, public policy-makers, and funding agencies to invest time and resources to ensure that effective active learning is incorporated into post-secondary mathematics classrooms" (p. 1) The *Guidelines for Assessment and Instruction in Statistics Education College Report* further stresses the importance of active learning in the mathematical sciences, encouraging instructors to use active learning to foster and enhance students' understanding and communication of statistics (Carver et al., 2016).

Despite these calls for instructors to adopt active learning teaching techniques, the use of such techniques in classrooms remains sparse (Apkarian et al., 2021). Further, active learning is challenging for GSIs to effectively implement because their knowledge about and use of active learning as instructors may be extremely limited (Justice, 2020). To learn how to best support GSIs in implementing active learning, our research investigates the evolution of and relationships between GSIs' knowledge about, emotions towards, and uses of active learning in the classroom.

Knowledge About Active Learning

An identified challenge in using active learning is having sufficient knowledge about the teaching practice (Finelli et al., 2014). When using active learning techniques, instructors draw upon several types of knowledge, including pedagogical content knowledge and generalizable content knowledge (Andrews, Auerbach, & Grant, 2019). However, many new GSIs have little teaching experience to help develop these types of knowledge and, when teaching, typically refer to previous experiences as a student which were largely lecture-based (Harmin & Melanie, 2006; Justice, 2020; Stacy, 2000). Thus, GSIs often have minimal understanding of what active learning is, or how active learning is used as an instructor.

Knowledge about active learning is an important factor to study in relation to the use of active learning in the classroom. Teaching experience is crucial but does not solely manifest into teaching knowledge for active learning (Andrews et al., 2019). Literature needs to expand through more complex study designs (e.g., longitudinal designs) to investigate how knowledge about active learning evolves over time and how to best support GSIs' early growth in knowledge about active learning.

Emotions Towards Active Learning

With the increased interest in educational research on emotions, researchers have started to identify the role of emotions in education (e.g., Boler 1999; Nias 1996). Currently, research on emotions has primarily focused on the relationship between emotions and instructors' work, development, and livelihood (Hargreaves, 2005; Zembylas, 2005). It is suggested that when an instructor is experiencing negative emotions, they may experience higher levels of burnout and job dissatisfaction (Keller, Chang, Becker, Goetz, & Frenzel, 2014).

With negative emotions identified as a potential challenge for teacher effectiveness and engagement in behavior, such as using active learning, it is crucial that we study how their emotions change over time to best understand how to develop and support GSIs' positive emotions towards active learning (Tyng, Amin, Saad, & Malik, 2017). Currently, there are no known studies that longitudinally investigate GSIs' emotions towards active learning, how their emotions evolve and relate to their knowledge about active learning, and how their emotions may impact their use of active learning.

Use of Active Learning

GSIs typically approach teaching based on their prior experiences in school as a student (Justice, 2020). Further, when using active learning, new instructors often implement less complex techniques that take less time to implement (e.g., minute papers), often adapting rather than adopting practices and unknowingly compromising their effectiveness (Vickrey, Rosploch, Rahmanian, Pilarz, & Stains, 2017). The two challenges discussed above (knowledge about and emotions towards active learning) may play a role in limiting new instructors' use of active learning and its effectiveness in student learning.

Auerbach et. Al (2018) compared expert and new instructors' knowledge about active learning by studying how each group of instructors held students accountable, whether the instructor elicited and responded to student thinking, and opportunities afforded to students to generate their own ideas and work. Noticeable differences were observed between new and more experienced instructors, with the newer instructors considering less about how their techniques hold students accountable or promote student thinking. Additionally, emotions and their role in influencing actions and behaviors (like using active learning) has taken its place in theories of action (Zhu & Thagard, 2002). Experiencing negative emotions may inhibit the quality of teaching and the use of active learning.

Our research about GSIs' longitudinal progression through three challenges—knowledge about, emotions towards and uses of active learning—will provide insight into how and when GSIs experience breakthroughs with these challenges. This information will offer insight into how we can best support GSIs to break through these challenges early in their teaching career and effectively use active learning in their classrooms.

Conceptual Framework

In this study, we frame GSIs' knowledge about, emotions towards, and uses of active learning under the assumption that "educators are learners," continuously learning through their teaching experiences. Transformative learning theory describes the process of transforming and updating habits of mind and points of view that shape an instructor's experiences (Mezirow, 1991). Mezirow (2003) defines transformative learning as: "learning that transforms problematic frames of reference, sets of fixed assumptions, and expectations (habits of mind, meaning perspectives, mindsets) to make them more inclusive, discriminating, open, reflective, and emotionally able to change" (p. 58). We posit knowledge and emotions, as frames of reference or habitual ways of thinking and feeling (Mezirow, 1997), shape how a GSI uses active learning when teaching. For example, a GSI may know that using active learning helps students' retention of the material which influences more frequent use of active learning teaching techniques in their classroom.

Instructors can transform their frames of reference through their teaching experiences and professional development, and this type of transformation only occurs when learning is meaningful (Mezirow, 1997). During meaningful learning, educators critically reflect upon new knowledge gained during their individual and collective learning experiences and relate it to their pre-existing knowledge (Ausubel, 1968). This reflection reshapes how they interpret and engage with future experiences, modifying their existing frames of reference. We believe this continuous process of transformative learning, as illustrated in Figure 2.1, portrays how we can study and understand GSIs' uses of active learning over time. By continuously reflecting upon what is being learned and what is already known, GSIs' experiences with using active learning during instruction continues to evolve, shaping their knowledge about, emotions towards, and ultimately their uses of such teaching practices. For example, a GSI may reflect on feedback about their uses of active learning through student comments and compare it to what they already know about teaching. This reflection may change the knowledge a GSI possesses and their emotions towards active learning, updating their frames of reference for their next reflection process, and influencing teaching practices moving forward. This research specifically investigates knowledge and emotions as frames of reference towards the use of active learning to better understand GSIs' evolution over time, their relationship with the use of active learning, and their relationship with each other.

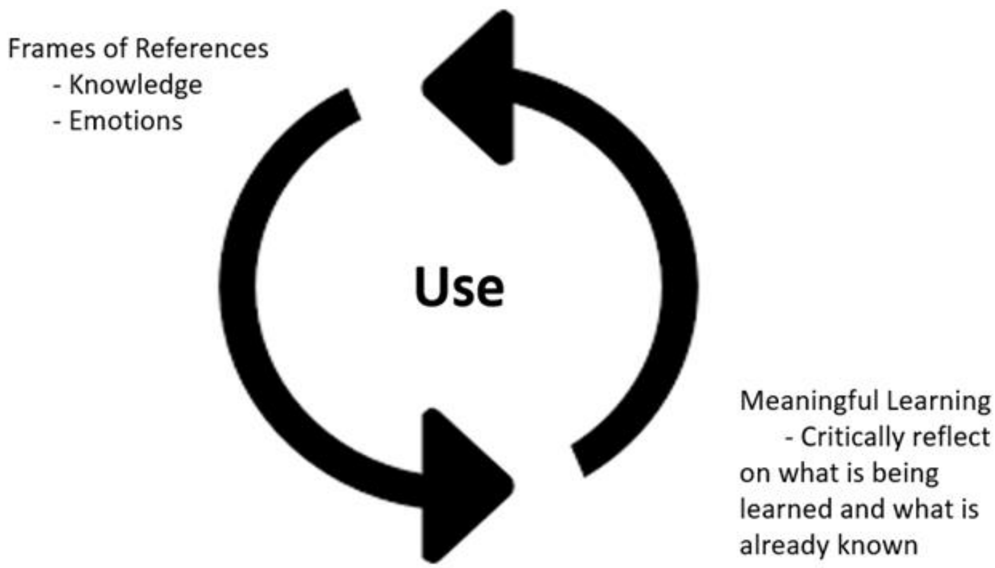


Figure 2.1: Framework describing how GSIs experience active learning over time.

Methods

In this study, we used a collective case study approach, where each case is defined to be each GSI. This approach encompasses both a within-case and a cross-case analysis. A within-case analysis helps describe and explain a GSI's longitudinal progression of their knowledge about, emotions towards, and uses of active learning. Because GSIs' knowledge about, emotions towards, and uses of active learning can differ across individuals, we also employed a cross-case analysis to explore processes and outcomes across each case. The cross-case analysis helps develop more sophisticated and powerful descriptions of GSIs' knowledge about, emotions towards, and uses of active learning within the contextual boundary of a mathematical sciences department at a land-grant research university in the Rocky Mountain region. At this university, a GSI is typically the lead instructor of one section of a freshman- or sophomore-level mathematics or statistics course each semester. These courses often include College Algebra, Pre-Calculus, Calculus I, and Introductory Statistics, which have approximately 30–40 students per section.

Department Support for GSIs

All new GSIs in the mathematical sciences department attended a week-long orientation prior to the start of the academic year. At the beginning of this research study in Fall 2017, the orientation included workshops and activities focused on building a community of instructors, preparing GSIs for their first week of teaching, and encouraging the use of active learning. This orientation offered GSIs multiple opportunities to interact with their peers and experienced instructors while learning about their upcoming roles and responsibilities as GSIs.

During the academic year, the new GSIs continued to meet weekly in Fall 2017 and monthly in Spring 2018, attending interactive one-hour workshops that exposed them to innovative teaching methods, techniques, and tools (Appendix A). Each workshop was structured so that GSIs could observe, explore, and reflect upon a pre-determined topic. We note that the first three authors co-developed and co-led the GSI development program.

In addition to the development program for new GSIs, all GSIs in the mathematical sciences department had weekly meetings with their Student Success Coordinator (SSC). Each

SSC is responsible for coordinating the instruction and assessment of all sections of an undergraduate course. During weekly course meetings, the SSCs discussed the material GSIs would teach for the week, encouraging them to use active learning techniques.

<u>Participants</u>

At the beginning and the end of the first semester of the GSI development program in Fall 2017, we surveyed all new GSIs (*n*=20) about their perceived roles as educators and their definitions of active learning. Because an instructor's perceived role can help shape their experiences with active learning (Grasha, 1994), we chose to group GSIs based on how they perceived their roles as instructors at the beginning and the end of their first semester of teaching. Based on the GSIs' written responses, we identified two different roles: facilitator and lecturer. GSIs in the facilitator group wrote about the personal nature of instructor-student interactions, and the value of students making informed choices and self-discovering the value in the mathematical or statistical content. GSIs in the lecturer group noted a one-directional flow of knowledge, stating they were responsible for passing information onto their students. We purposefully selected two GSIs (Table 2.1) who were representative of GSIs who we identified as either a lecturer or facilitator.

Max (pseudonym) is a self-identified female from the United States who consistently described herself as a *lecturer*: "My role as a teacher is to assist students in learning through lectures, worksheets, homework, and doing my best to answer questions that arise." Andy (pseudonym) is a self-identified male and international student who consistently described himself as a *facilitator*: "I function majorly as a facilitator of learning. I am also a learner in my own class." Max had no prior teaching experience and "had never heard of active learning before." During the week-long orientation, Max wrote that she planned to use "group work, smaller lectures before group work, [and] different colored pens to emphasize stuff." In contrast, Andy had experience in a K-12 setting from 2012 to 2017 and had completed a professional development program for instructors at another institution in 2016. At the start of the semester, he noted that he planned to use "grouping, worksheets, pair and group teaching, differentiated learning, [and] formative assessment" in his classroom to actively engage students. These differing perceptions of their role as an instructor allow investigation into when breakthroughs with active learning occur for each type of instructor.

	Max	**Andy**
Degree Program	M.S., Mathematics	Ph.D., Mathematics Education
Perceived Role as Instructor (Fall 2017)	Lecturer	Facilitator
Demographics	Female, United States	Male, International
Prior Teaching Experience	None	Completed a different teacher development program; Lead instructor of mathematics classes
Initial Understanding of Active Learning (Fall 2017)	"I had never heard of active learning before."	"Engage our students in the learning process."
Courses Taught (Fall 2017-Fall 2019)	Survey and Applications Course in Calculus (4 sections) College Algebra (2 sections) Introductory Statistics (1 section)	Survey and Applications Course in Calculus (2 sections) Calculus 1 (4 sections)

Table 2.1: Descriptions of the study's participants.

Data Collection

This research used similar data collection techniques during each semester, as shown in Table 2.2. We used interview and survey data to provide a thorough description of GSIs' knowledge about, emotions towards, and use of active learning during each semester of the study, and these data sources helped triangulate what we observed in their classrooms through video recordings.

	Fall 2017	**Spring 2018**	**Fall 2018**	**Fall 2019**
Data Collected	Pre- and post-semester surveys	Classroom observation Pre- and post-observation interviews	Classroom observation Post-observation interview Survey	Interview Survey

Table 2.2: Description of data collection.

From Spring 2018 through Fall 2019, each participant completed four semi-structured, recorded interviews. They answered questions about their knowledge about, emotions towards, and uses of active learning in their classroom (Appendix B). Sample questions included: *What is your definition of active learning? When active learning was introduced, how did you feel about it? If you were presented with a new active learning technique, how comfortable would you be trying to implement it in your class?* Before each interview, GSIs were reminded their identities and recorded data would remain confidential. Each GSI was reassured that the data collected would not have any impact on their employment as a GSI was encouraged to answer each question truthfully to the best of their ability.

Each participant also completed four free-response surveys (Appendix C). At the beginning and the end of Fall 2017, the surveys asked about their perceived roles as educators and their definitions of active learning. The other two surveys inquired about how each GSI thought and felt about active learning, how they used active learning techniques in their classrooms, and how (if at all) their knowledge about, emotions towards, and use of active learning changed from previous semesters. The last survey given in Fall 2019 also asked about the challenges GSIs faced when using active learning. Some questions included: *How do you define active learning? Has your definition changed since last semester? When you first discovered active learning, what challenge(s) did you face when trying to implement such techniques?*

Each participant agreed to be video recorded while teaching on two separate occasions, once during Spring 2018 and once during Fall 2018. Max was recorded instructing a Survey and Applications Course in Calculus both semesters, and Andy was first recorded instructing a Survey and Applications Course in Calculus and then Calculus I. To capture how the participants used active learning, the video recordings occurred on days when each GSI self-reported that they planned to use active learning techniques; each GSI was given multiple weeks to select a date. The researchers were not present during the filming of the class to ensure each GSI didn’t feel undue pressure when teaching.

Data Analysis

The first author wrote memos of important ideas and responses for each survey transcribed and during each interview to initially explore these data. Memos include descriptions of the participant's tone, body language, and other verbal and non-verbal cues when answering questions. Additional memos were created when transcribing each interview to build upon and provide evidence of validity in the researcher's memos. We used an observation protocol (Appendix D) to identify how GSIs were using active learning. The following sections detail the analysis of these data regarding each GSI's knowledge about, emotions towards, and uses of active learning.

Knowledge

To investigate GSIs' knowledge about active learning, we utilized interview transcripts, survey data, and memos as our primary data sources. We used an iterative coding process by repeating the data analysis process to provide evidence of reliability towards the formation of the codes. Through this iterative process, the lead author used a provisional coding approach, beginning with a list of researcher-generated codes found in the literature on graduate students, active learning, and teaching experiences (Miles, Huberman, & Saldana, 2014); for example, *student involvement* is a term used when defining active learning (Bonwell & Eison, 1991). This list of generated codes and their working definitions created our initial codebook (Table 2.3). Other codes emerged throughout the study based on the GSIs' responses to the interview and survey questions.

Code	Definition
Cognitive engagement	Psychological effort students devote to interacting with others and the material (Bonwell & Eison, 1991)
Student involvement	Students' willingness to take on the learning task (Freeman et al., 2014)
Learning process	Process of acquiring new understanding, knowledge, behaviors, or skills for oneself (Freeman et al., 2014)
Pause for reflection, Think-Pair-Share, Large-group discussion, etc…	List of active learning techniques (Center for Research on Learning and Teaching, n.d.)
Engage, Explore, Explain, Elaborate, Evaluate	5 E Instructional Model: Five phase sequence that instructors facilitate to put students at the center of learning (Kudryashova & Rybushkina, 2016)

Table 2.3: Descriptive codes from literature.

At each timepoint, we combined similar codes within a single case to create initial themes, which we then discussed and refined as a research team. For example, Table 2.4 provides codes and their descriptions that were generated based on Max's responses to an interview question asked in Spring 2018; the codes including *detail*, *student involvement*, and *student attitudes* all describe a general theme of *positive engagement.*

Participant: Max – Spring 2018	Code	Code Description/Evidence	Theme
Question: How do you define active learning?	Detail	Max talks about how active learning lets students provide more detail to an answer, instead of just a one-word answer.	Positive Engagement – showing a willingness to remain cognitively positive and effort to be immersed with others in the content.
	Student Involvement	Max discusses how she thinks active learning is when students are not just sitting and watching the teacher do problems. They are willing to be involved in the lesson.	
	Student Attitudes	Max relates active learning to the idea that students should not put themselves down when doing mathematics.	

Table 2.4: Coding example for one question for Max (Spring 2018).

Emotion

To investigate emotions about active learning, we used interview transcripts, survey data, and memos as our primary data sources. For each interview and survey, we generated codes from the data using a similar iterative approach that was used in analyzing knowledge about active learning.

Once we created initial codes from iterative evidence across the data, we implemented a deductive approach to group similar codes together. Our deductive approach used pre-determined emotions from Plutchik's (2001) Wheel of Emotion (Figure 2.2) as descriptive themes of the data-generated codes. This wheel suggests eight primary emotion dimensions represented by sectors (e.g., love, contempt), and Plutchik divided each primary emotion into different subgroups. Intensity of emotion is characterized by its position on the wheel; emotions closer to the center describe more intense emotion, and emotion polarities are observed through opposite ends of the wheel.

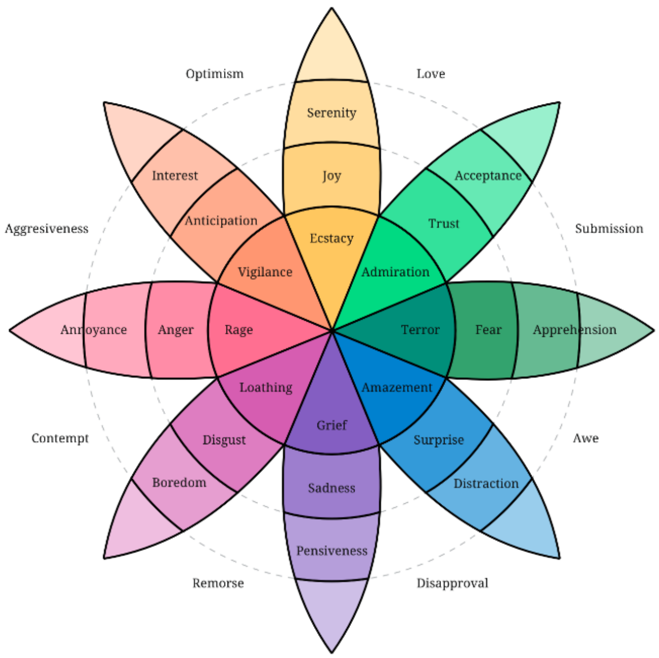


Figure 2.2: Plutchik's (2001) Wheel of Emotion that uses a color wheel to highlight different types of emotions. Emotions closer to the center of the wheel are considered more intense. Adapted from "The Nature of Emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice," by R. Plutchik, 2001, *American Scientist*, *29*(4), p. 349. Copyright Creative Commons Attribution License.

As an example, Andy shared that prior to our study, he thought active learning would never work and its use was "bullshit." We coded Andy's statement with the initial codes, *frustration* and *irritation*, which were then summarized by Plutchik's Wheel of Emotion as *anger*.

Use

To understand how GSIs implemented active learning techniques in their classrooms, we analyzed their self-recorded classroom teaching videos. Two of the authors analyzed the video data using an observation protocol created by the research team. Referencing active learning definitions, guidelines, and national recommendations, the two authors piloted the protocol on classroom teaching videos from three GSIs who were not selected to participate in this research study. During the pilot, the two authors independently coded each video, recording the time stamp(s) when active learning was observed, and identifying the type of active learning technique(s) used (Center for Research on Learning and Teaching, n.d.). After discussing their results, the two authors modified the protocol (with approval from the research team) to also include rich written descriptions of how the GSIs implemented active learning and an additional section to include visual cues of student engagement.

Using the modified protocol, the same two authors first independently analyzed Max's and Andy's classroom teaching videos and then later corroborated all identified instances of active learning with one another. For example, one author wrote, "Max first reviewed a u-substitution integral problem on the board with the entire class. She asked students questions as she worked through the problem on the board (~ 4-minute mark)." These identifications and descriptions of each GSI's use of active learning across timepoints were compared among the two authors to reach consensus. Additionally, we asked interview questions targeting how each GSI uses active learning to triangulate these observations from the classroom video recordings.

Breakthroughs

When and how breakthroughs with active learning occurred for each GSI are discussed below through an interpretivist thematic comparison of content of the themes across each timepoint (beginning Spring 2018, end of Spring 2018, Fall 2018, Fall 2019). During comparisons, we noted if themes within each of these constructs showed evidence of a change in content across time and assessed whether a change represented an increase in knowledge about, more positive emotions towards, or more frequent uses of active learning. A positive shift in themes across timepoints was labeled as a breakthrough. Our conceptual framework outlines each GSI's process on how these breakthroughs may occur between timepoints. Potential breakthroughs occur when experiences, such as teaching or engaging in conversations, incite critical reflection about what is known and what is learned about active learning.

We also present little to no evidence of breakthroughs to better contextualize when breakthroughs occurred. We note and discuss a shift to a lesser understanding, more negative emotions towards, or less uses of active learning. Further, we also compare the themes describing each GSI's knowledge about, emotions towards, and uses of active learning across the two cases within each of the four timepoints.

Results About Max

Coming into this study, Max had no prior teaching experience, nor had she heard of active learning. Below, we outline Max's breakthroughs with active learning in the context of her knowledge about, emotions towards, and uses of active learning in her teaching (Figure 2.3).

Time	Knowledge	Emotion	Use
Beginning Spring 2018	Understood as getting your students *involved*	Communicated emotions of *apprehension*	*Passive* usage of active learning
End Spring 2018	Better articulated her knowledge of active learning by understanding it as *positive engagement*	Communicated emotions of *apprehension* and *fear* stemming from how students would respond	Despite wanting to use active learning, usage remained *limited*
Fall 2018	Description remained consistent, centered around *engagement*	Described emotions of *trust* and *acceptance*	Started at the beginning of the semester, primarily using *group work*
Fall 2019	Description remained consistent, centered around *engagement*	*Apprehensive* of techniques that required students to work outside of class but showed *trust* in in-class techniques.	Continued to use active learning by *asking questions* and *letting students work together*

Figure 2.3: Max's evolution and breakthroughs with active learning. Breakthroughs are noted with blue arrows.

<u>Spring 2018 Breakthrough (1)</u>

Max's definitions and descriptions of active learning evolved throughout Spring 2018. At the beginning of Spring 2018, Max described active learning as students doing something more than just sitting and listening to the instructor: "[active learning is] getting your students involved in class, and not just...sitting there watching you lecture to them." Max explained, "I think it's [active learning] the difference between a one-word answer and more of an extensive answer, and it's about involvement of your students." She thought that active learning techniques were helpful to get students to talk to each other and found learning about these techniques rewarding for her overall growth as an instructor.

At the end of Spring 2018, between her second and third semester teaching, Max showed a more advanced understanding about active learning by elaborating more on what she meant when she described active learning as students not just sitting. For Max, not sitting meant that students were not passively learning the material but instead learning through "some form of active engagement." She articulated a more advanced understanding of active learning by discussing it further throughout the lens of engagement. When explaining engagement, she detailed students being cognitively engaged by "really paying attention" and understanding the transitional steps when learning proofs. She thought active learning teaching techniques were helpful for "get[ting] these students involved with one another... [and to] keep students awake." Max largely attributed her more advanced understanding of active learning (i.e., breakthrough) to conversations with other instructors and the GSI development program she attended in Fall 2017.

We did not observe Max experience a breakthrough in her emotions towards and her use of active learning during Spring 2018. Max's emotions towards active learning were strongly and consistently influenced by her own experiences as a student. She had positive experiences learning in lecture classes stating, "As far as [my] classroom experience, for the most part, it has been very, very positive. My math classes in undergrad were very much a lecture-based thing." When reflecting on her experiences with active learning, she said, "I know for me when I was a student and forced into those situations I didn't prefer having to work in groups." In fact, as a student, Max "fully avoided having one [flipped (non-lecture-based) class] because I didn't think that was how I learned." Thus, as an instructor, Max relayed emotions of apprehension and fear of using active learning. This related to concerns about how students may interpret her use of active learning. Max explained that using different active learning techniques may result in students becoming adversarial towards each other or make her students feel like "children."

During this time, Max was also hesitant to use active learning in her classroom and questioned how it could work in a mathematics class, stating, "That's awesome, but how in the heck do you do that [active learning] for math?" Max was unsure about active learning, wrestling with conflicting views about the benefits students may receive and her own learning preferences as a student. She reflected: "I think for the most part it benefits them in the sense that it challenges them to understand the material enough to explain it to another person on a pure level. ... [However,] I was a student that hated group work so I kind of battle between what I like and what I don't." Video data depicted Max predominantly lecturing to her students and then assigning a worksheet midway through the class that students primarily completed individually. Despite a more advanced understanding of active learning, Max's emotions towards active learning remained largely negative and her use of active learning was extremely limited.

Fall 2018 Breakthroughs (2 & 3)

During her third semester teaching in Fall 2018, Max experienced a breakthrough in both her emotions towards and use of active learning, shifting to align with her more advanced knowledge about active learning. She displayed trust and acceptance and conveyed optimism when discussing how active learning can help improve student understanding. She felt excited about using active learning, stating that she was "really enjoying it" and found it "fun to see your class interacting with one another."

As her emotions towards active learning evolved positively, Max's use of active learning was more apparent. In the classroom video data, we observed Max ask more questions during class and facilitate conversation between students, frequently encouraging students to discuss with their group members. For example, Max asked the following questions which engaged multiple students in conversation: “Did you guys converse about this?” and “Do you understand their explanation?”

We observed Max experience breakthroughs in both her emotions towards and her use of active learning in Fall 2018, the same semester when she first expressed a desire to seize the opportunity to try and use active learning in her classroom at the beginning of the semester. Thus, Max’s breakthrough in use of active learning was observed at the same time she was both knowledgeable about active learning and experienced a breakthrough in emotions towards active learning. Max explained that this alignment between her knowledge about, emotions towards, and use of active learning led her to being “a lot more open to doing [active leaning] as much as possible instead of being like, no, I like to lecture and don’t really want to let go of this.”

Fall 2019

In Fall 2019, we observed Max initially re-experience some negative emotions towards active learning. During this semester, her frames of reference were challenged by being assigned to teach in a flipped classroom where students completed course content readings outside of class time and spent class time working through problems. Teaching within this different classroom structure, Max expressed hesitation, annoyance, and apprehension with active learning due to teaching in an unfamiliar classroom setup that she often avoided as a student. She substantiated her annoyance and apprehension, explaining that she’s a “control freak.” She did not want to use active learning strategies that relied on students’ preparation before class, because she could never be sure they were actually prepared. Over time, she started to become more comfortable: “You kind of take yourself and your personal feelings about it [active learning] out of the equation as much as you can so that you can do your job as a teacher as it is stipulated by your boss. Which can be frustrating, and I think initially it was frustrating and hard for me. And now it's going much better."

Despite conveying annoyance and apprehension towards the structure of active learning in a flipped classroom, she eventually expressed trust in active learning's ability to help students learn and eventually suggested that using active learning was easier after gaining teaching experience in a flipped classroom because she did not have to figure out how to implement group work. When asked if she would want to try new active learning techniques, Max discussed her apprehension, saying she was uncomfortable implementing new techniques because she was not sure how those would fit in with the structure of her class.

Throughout the study, when Max reflected on new information about active learning, she compared it to what she experienced as a student and her own classroom practice. These more traditional experiences that made up Max's initial frames of reference were reinforced by the teaching techniques she used in subsequent semesters, suggesting that simply increasing knowledge about active learning may not be enough to manifest its use in the classroom. Max was using "what worked" for her as a student when teaching until she subsequently experienced breakthroughs in her emotions towards and her use of active learning in Fall 2018.

Later, when Max became apprehensive about using active learning in a flipped classroom, her new, more positive frames of reference about active learning may have helped her to not revert to only traditional means of teaching: "There has to be a middle ground of these two [techniques]." Unlike previous semesters teaching, Max showed no fear in how active learning would be perceived by students: "They don't necessarily enjoy that I [use active learning], but I think it really gets them to have to think about it and synthesize some knowledge."

Results About Andy

Andy, who came into the study with prior teaching experience and an understanding of active learning, showed evidence of experiencing breakthroughs in emotions about and use of active learning, both in similar and different manners than Max. Below, we outline when and how Andy's breakthroughs with active learning occurred (Figure 2.4).

Time	Knowledge	Emotion	Use
Before Study	Attended professional development which taught active learning as *engagement*	Communicated emotions of *apprehension* and *anger* when introduced to active learning	*Forced* to use active learning; predominantly used *group work*
Beginning Spring 2018	Described active learning, at its core, as *engaging* students	Shared emotions of *trust* and *acceptance*	*Facilitated discussions* and *activities* with students
End Spring 2018	Continued to describe active learning as *engaging students*	Emotions, remained consistent sharing his *trust* and *acceptance*	*Flexible* in using different types of active learning; Took on more of an *interest* and became curious on how research and recommendations would play out
Fall 2018	Continued to describe active learning as *engaging students*	Expressed *joy* and *admiration*	Remained *committed* to a *flexible* active learning approach
Fall 2019	Described active learning as students taking *active roles* in their learning	Expressed *joy* and *admiration*	Remained *committed* to a *flexible* active learning approach

Figure 2.4: Andy's evolution and breakthroughs with active learning. Breakthroughs are noted with blue arrows.

Before the Study

Unlike Max, Andy had prior teaching experience, so we establish his baseline knowledge about, emotions towards, and uses of active learning in this section. Andy explained that when he first started to teach in 2012, his knowledge about active learning increased through the professional development (PD) he took at his previous institution, but he initially thought "[active learning] would never work" and was apprehensive and angry about the idea of using active learning. During his initial PD experience, Andy "was exposed to this collaboration and active learning so...[he] had the definition of active learning, of what active learning should be."

Andy met active learning with frames of reference that were built upon more traditional experiences as a student and instructor. Andy had never experienced active learning as an undergraduate student; his undergraduate education was primarily traditional, fueling his apprehension and anger toward the idea of using active learning: "When [active learning] was [first] presented I was like this is bullshit. I never had the experience. I was like this would never work. How would you have students sit in groups? They will be making noise." He explained that before the PD, he used to be a dictator in his class:

> I used to be the almighty teacher. The one with the almighty knowledge. You have to pour everything out to the students. And you would see me sweating profusely because you know right from the outset, I wanted my students to learn mathematics.

However, while still learning about active learning, Andy was required to teach in a flipped classroom that required the use of active learning: "The school has the students sit in groups, so as the teacher you can't say, 'I want a traditional way.'" At that time, he questioned the idea:

“Students would be allowed to talk in a math class? No, that never happens.” By the end of his initial PD experience, he thought, “Oh! This is cool.”

Spring 2018 Breakthrough (1)

Andy, entering our study with an advanced knowledge about, positive emotions towards, and a consistent use of active learning, experienced a breakthrough in how he used active learning by the end of Spring 2018. At the beginning of Spring 2018, Andy described active learning as a tool to engage his students in learning the material, where learners are learning from each other. He detailed how using active learning (using group work and group discussions) helps students better learn, portraying trust and acceptance of active learning because its use “makes students to be proactive in their reasoning.”

By the end of Spring 2018, Andy demonstrated a more flexible approach to using active learning by implementing a variety of active learning techniques. When using active learning in his class, Andy started by sharing the day’s learning objectives, emphasizing the process of problem solving by telling students to “exercise their minds.” When facilitating classroom discussions, he questioned students’ logic and reasoning and focused on interpretations over correct answers. When administering a warm-up worksheet of five true or false questions, he asked students to work together and “think and be sure you know why this is true or false.” When his students asked him if the worksheet would be graded, Andy stated, “No. We just want to see what is happening!” When a student tried to confirm a true or false answer, Andy asked, “If it is true, what makes it true?” In short, Andy used a variety of active learning techniques to engage students with the material throughout the entire class; he asked groups to share and compare answers, held class discussion on the true-false warm-up questions, asked students to raise their hands if they agreed with an answer, and had a student show and explain their answer to the class. Andy also talked to students at eye level, seemingly making them feel more comfortable to communicate with him as he checked in with groups.

During this time, Andy said he was curious to try many active learning techniques because every class is different, and you need to try different techniques to find out what works. He also explained why he enjoys doing this:

> As a Ph.D. student I think I do more of thinking like, ‘Okay if I want to do research, what can I actually find out?’ So sometimes I think of all these things and say, ‘Okay let me use this thing in the class. Will it make sense? Will it work?

Similar to Max, having a sophisticated knowledge about active learning and emotions towards active learning that aligned with his knowledge may have helped Andy better manifest a targeted, consistent, and flexible approach to using active learning in the classroom.

Fall 2018 Breakthrough (2)

In Fall 2018, Andy detailed a breakthrough with his emotions towards active learning. Remaining consistent in his knowledge about active learning, Andy now saw instruction as synonymous with active learning, continuing to display admiration, and now ecstasy, towards active learning. He stated:

> I can say now that I feel uncomfortable seeing my students sitting individually. If they are not working in groups, it seems like I’m doing something wrong. You know at a point in time you come to this realization that this thing is now part of me.

He discussed some of the hesitations students might have about engaging in active learning, but asserted, “People might not think this, but I love that it’s all about the knowledge.”

Fall 2019

In Fall 2019, Andy’s knowledge about, emotions towards, and use of active learning remained unchanged as he continued to show an advanced knowledge about, positive emotions towards, and flexible uses of active learning in his classroom. During this time, he reiterated his role as a facilitator when defining active learning: “There is one thing I know about active learning for me… placing myself as a facilitator in the classroom whereby my strategies allow my students to...take active roles in the classroom.” He emphasized that the active roles students take include discussing with each other and solving problems

Although initially resisting active learning based on what his initial frames of reference were about teaching, Andy’s critical reflection on his experiences using active learning in his classroom facilitated updates in his knowledge and emotional frames of reference. During our study, as Andy gained more experience with active learning, his frames of reference surrounding his knowledge and emotions were reinforced, setting the stage for a committed and flexible approach to using active learning.

Cross Case Comparison

We noted similarities and differences about how and when Andy and Max experienced breakthroughs during our study. We observed Max experience a breakthrough with knowledge, as she started to describe in detail how she could get her students engaged in the material and with others. Andy shared a similar understanding of active learning upon entering our study. Both attributed their knowledge about active learning, in part, to their initial experiences in PD programs. Once understanding active learning through the lens of engagement, each GSI remained consistent in how they understood active learning for the duration of the study. Further, each GSI experienced a breakthrough in having more positive emotions towards active learning. Once experiencing a breakthrough in emotions, we observed an increased willingness (in each GSI) to use active learning techniques in their classrooms, and both GSIs shared how they trusted the use of active learning to heighten student learning.

Overall, Andy shared more consistent and extremely positive emotions than Max during the duration of the study. Max showed evidence of reservation towards active learning after her breakthrough, largely attributing this to the new flipped classroom structure and additional responsibility her curricula was putting on her students. Andy, initially introduced to active learning through a flipped classroom, shared that the class structure helped him understand how he could interact with students and not be the “grand teacher” in front of his students.

Discussion

Even though national recommendations suggest the use of active learning techniques in undergraduate statistics and mathematics instruction, many instructors rely on lecture-based instruction, believing that their students learn best in that environment (Johnson, 2019). We observed this reliance on lecture-based instruction with Max and Andy, particularly as they detailed their first-time teaching. Literature on beginning instructors suggests that the first year of teaching is a survival year (Beisiegel, 2019). During this stage, the instructor’s main concern is whether they can survive, reflecting on questions such as, '"Can I get through the day in one piece;' 'Can I really do this kind of work day after day?'" (Katz, 1972, p. 3). While in survival mode, instructors are first learning about teaching and are not necessarily ready to enact more demanding instructional approaches, such as active learning (Beisiegel, 2019; Katz, 1972). However, with the increasingly important role GSIs play in undergraduate mathematics and statistics education, along with updated best practices in teaching, it's critical GSIs implement teaching strategies that align with current research-based recommendations on how to effectively teach. Our investigation into when and how GSIs break through with active learning promotes discussion about how to support GSIs' use of active learning.

GSIs need to feel confident in and be challenged by their course structure before implementing active learning techniques. Andy's and Max's use of active learning were heavily influenced by the structure of their courses. Andy's breakthroughs with active learning were kickstarted during his first semester teaching when he got to experience active learning in a flipped classroom that required its use; during this experience, he gradually became more confident in the use of active learning. Max, whose first teaching experiences were in more traditional lecture-style classrooms, didn't experience a breakthrough in her emotions or use of active learning until her third semester teaching when she made the choice to start using active learning right away at the beginning of the semester. This finding may support research suggesting that instructors have an easier time using active learning when the curriculum is inquiry-based and necessitates the use of active learning for instruction from the beginning (e.g., Haack, 2008). Having a curriculum and classroom environment that supported active learning through the required use of group work allowed Andy to expand his knowledge about, and update his emotions towards, active learning. Initially resisting active learning, curricula that supported his use of active learning allowed him to conclude, “Oh! This is cool.”

Max, who did not know about active learning prior to our study, struggled to break through in using active learning until she had more teaching experience. It wasn’t until her third semester teaching (Fall 2018) that she was able to challenge her more traditional styles of instruction, citing that she either didn’t know enough about active learning or “didn’t necessarily know how to implement [active learning] in an effective way especially because [she] didn’t start right off the bat doing it.” Starting the Fall 2018 semester, Max admitted that she was still hesitant to use active learning but wanted to challenge herself to teach outside of her comfort zone because she never “ever wanted to be a stale teacher” and wanted her students to talk to one another.

Unlike Andy, Max did not have a curriculum to help challenge her use of “what works” when she first started teaching and instead had to challenge herself to incorporate the use of active learning into her classroom despite still having doubts. A lack of curriculum and classroom support to help Max implement active learning techniques in her classroom may be one of the reasons why her knowledge about and emotions towards active learning didn’t break through and align until her third semester teaching. For Max, starting at the beginning of the semester helped influence more consistent use of group work when teaching. In Fall 2019, when in a flipped classroom, Max articulated that using group work was easier than before.

Referencing "using what works" for her as a student when teaching provides evidence that the roles of knowledge about and emotions towards active learning may drive a feedback loop that promotes or inhibits GSIs' breakthroughs in using active learning. Before the Fall 2018 semester, Max had made minimal attempts to use active learning, so her knowledge about and emotions towards active learning had solely arisen from what she had learned during the GSI development program and not through prior experience as a student or teacher. Reflecting on her experiences as an undergraduate student, she started teaching the way she had been taught before. This created more traditional teaching experiences as a GSI in her initial semesters of teaching, stalling breakthroughs with the use of active learning until her third semester. Max used and was reinforced by traditional lecture techniques that did not create a learning environment where the students could become adversarial or feel like children, and ensured the students were getting all the information they needed from her—all deterrents for Max to use active learning. These experiences also may have inhibited the use of more complex active learning techniques when she started to use them in Fall 2018, which predominantly consisted of group work. For example, while teaching in a flipped classroom, Max explained that she was not comfortable introducing other active learning strategies besides group work. This highlights the importance of challenging the idea that novice instructors, such as GSIs, cannot use active learning in their beginning years of teaching. Evidence may suggest that the techniques GSIs initially use may influence and inform their subsequent semesters of teaching.

Additionally, as both GSIs solidified their knowledge about active learning, each GSI's emotions towards active learning may have helped fortify positive frames of reference that encouraged them to not revert to using more traditional measures of active learning. Research suggests that positive and negative emotions greatly influence individuals' knowledge, creation, learning, and memory while at the same time acting as a motivator for action and behavior (Kremer et al., 2019; Tyng et al., 2017).

When Max started to raise hesitations with active learning again in Fall 2019, she did not resort to predominantly traditional ways of teaching. A previous breakthrough in her emotions towards active learning updated her frames of reference, allowing her to reflect on newfound hesitations with trust and acceptance. Using these frames of reference to reflect on these new emotions, Max suggested, "There has to be a middle ground of these two [techniques]." Andy, with extremely positive frames of reference towards active learning, explained that, despite receiving some student resistance, he would always commit to a flexible active learning approach. This is similar to Kremer et al.'s (2019) findings about how emotions affect learning for 60 first-year internal medicine residents. They found that negative emotions towards the practice were associated with learning less about and ultimately less time engaging with the learning tasks. Possessed knowledge is a precondition for action (Funke, 2017) but does not initiate or guarantee the action. This evidence suggests that careful consideration into how PD can best support GSIs' emotions towards active learning may be necessary to best support GSIs' use of active learning in the classroom. We propose further investigation into how PD can best support GSIs' emotional intelligence and breakthroughs.

Lastly, the alignment of breakthroughs may be key to helping GSIs use active learning. Before using active learning, both Andy and Max first had breakthroughs that aligned their respective knowledge about and emotions towards using active learning, suggesting that supporting a GSI's knowledge about, as well as their positive emotions towards active learning may be important for promoting GSIs' early use of active learning. PD targeted only on increasing GSIs' knowledge about active learning may not result in their use of it. Each GSI in our study used active learning when they displayed both sufficient knowledge about and showed evidence of a breakthrough with positive emotions towards active learning that aligned with their knowledge. Max, despite understanding active learning through the lens of engagement, did not start using active learning in her classroom until she subsequently experienced a breakthrough in her emotions towards active learning.

Scope and Extensions

Although the above suggestions for PD were developed from existing literature and evidence from this research, further investigation into emotional intelligence's role in PD for GSIs needs be conducted. Emotional intelligence is commonly defined as the ability to perceive, use, understand, and manage emotions. Research has shown that emotional intelligence can be improved through training (Gilar, Pozo-Rico, Sánchez, & Castejon, 2019). Understanding what and how emotional intelligence competencies are addressed in a PD setting may help PD facilitators support GSIs in better aligning their more sophisticated understandings with more positive emotions towards active learning, promoting its use in beginning semesters.

When Max felt more confident and committed to using active learning, she disrupted her feedback loop by intentionally choosing to use active learning techniques at the beginning of her third semester teaching. She cited that this happened because she "did not want to become a stale instructor." When Andy was required to use active learning for the first time as an instructor (prior to this research study), his initial negative emotions towards active learning quickly changed after he saw the benefits play out in his classroom. Additional research should investigate other influential factors related to knowledge about and emotions towards active learning to further understand why these breakthroughs in active learning occur and how PD can support GSIs' breakthroughs with active learning. This understanding could help break the start of feedback loops that may influence more traditional styles of teaching.

Active learning materializes differently through more and less complex techniques in the classroom based on a GSI's knowledge about and emotions towards active learning (Bonwell & Eison, 1991). Future research should investigate both if and how GSIs make the transition to incorporating more complex active learning techniques in their classroom. Additionally, future research should collect data at more frequent timepoints within and across semesters to obtain higher resolution data for investigating when breakthroughs occur with active learning. It may be particularly useful to collect data at times within semester when GSIs critically reflect and update their frames of references, such as after they solicit feedback from students about their teaching or observe another instructor use active learning.

In this study, we narrowed the complexity of a GSI's frames of reference to their knowledge about and their emotions towards active learning. The frames of reference a GSI holds towards active learning are more complex and can be thought of as a set of criteria or stated values in relation to which GSIs use to make judgments. Expanding, or investigating different constructs, such as teaching beliefs and values, that make up frames of reference would lead to a better overall understanding of the complexity of GSIs using active learning in their classrooms.

Conclusion

The evidence from this research helps continue to build an argument that GSIs can use active learning early in their development with proper support and offers information about how to support GSIs in the future. Sufficient knowledge about active learning is foundational, but not sufficient for the use of active learning in a GSI's classroom. Emotions held towards active learning play a key role in how a GSI implements active learning, if at all. Because of this, PD needs to not only focus on GSIs' knowledge about active learning but also implement strategies to help GSIs understand and manage positive emotions towards active learning to increase and sustain its use in the classroom.

Further, PD programs need to offer ongoing support for GSIs' evolving needs as their knowledge about and emotions towards using active learning in the mathematical sciences change over time. Although knowledge about active learning consistently grew for each GSI, emotions towards active learning were more volatile for Max. Longitudinal studies, such as this one, are imperative to understand how PD programs can support the refinement of GSIs' instructional practices, helping to better manage positive emotions towards these techniques.

Acknowledgements

We thank Andy and Max for participating in this study and sharing their experiences with active learning.

A FRAMEWORK FOR UNDERSTANDING STATISTICS INSTRUCTORS' MOTIVATIONS TO USE ACTIVE LEARNING

Contribution of Authors and Co-Authors

Author: Elijah S. Meyer

Contributions: Selected frameworks, interpreted frameworks in new context, and wrote manuscript.

Co-Author: Jennifer L. Green

Contributions: Discussed frameworks and edited drafts of manuscripts.

Co-Author: Stacey A. Hancock

Contributions: Edited drafts of manuscripts.

Manuscript Information

Elijah S. Meyer & Jennifer L. Green & Stacey A. Hancock

To be determined.

Status of Manuscript:
X___ Prepared for submission to a peer-reviewed journal
____ Officially submitted to a peer-reviewed journal
____ Accepted by a peer-reviewed journal
____ Published in a peer-reviewed journal

Publisher: To be determined.

Abstract

National recommendations, such as the Guidelines for Assessment and Instruction in Statistics Education (GAISE) College Report, encourage the use of active learning within statistics classrooms to help promote student understanding and critical thinking. Despite these recommendations, instructors commonly use lecture-based teaching practices to teach statistics when faced with various challenges, including a limited understanding of active learning techniques. This gap between recommended and enacted instructional practices presents a need to better understand why statistics instructors use active learning to help inform future recommendations on how to support statistics instructors' use of such techniques. This article presents a framework for understanding why statistics instructors use active learning. Through the lens of self-determination theory (SDT), we discuss how these reasons, or motivations, characterize the use of active learning from least to most self-determined and describe the positive behavioral outcomes associated with more self-determined behavior. We highlight SDT's emphasis on one's needs for autonomy, competence, and relatedness as a tool for promoting more self-determined uses of active learning. Additionally, we connect these needs to self-regulation learning (SRL) theory, discussing how fulfillment of these needs can promote a statistics instructors' use of active learning. Lastly, we contrast key features of SDT with another commonly used motivational framework and highlight the benefits of using SDT to investigate statistics instructors' motivations to use active learning.

Introduction

Statistics is often defined as "the science of learning from data, and of measuring, controlling and communicating uncertainty" (Davidian & Louis Thomas, 2012, p. 12). In practice, statistics is used as a flexible and powerful tool across a variety of disciplines including business, academia, and government. Within these disciplines, statistics can act as a means for communication, an avenue for discoveries, and a tool to make data-informed decisions through the analysis of data. Moreso now than ever, it is widely recognized that statistics is among the most important quantitative subjects available through university curricula, as courses in statistics afford students the opportunity to find structure in and a deeper understanding of data (Weihs & Ickstadt, 2018).

To help ensure students develop this deeper understanding of data, national recommendations for teaching statistics encourage a shift from instructor-centered styles of lecture to more engaging, student-centered instructional practices, such as the use of active learning techniques (Carver et al., 2016). Active learning techniques are any classroom-based practices that engage students in their learning through means of problem solving or discussion, contrasting to more traditional lectures, where students passively receive information from the instructor (Nguyen, et al., 2021; Prince, 2013). Examples of active learning techniques include small-group collaboration, jigsaw discussions, and think-pair-share (Center for Research on Learning and Teaching, n.d.).

Research has shown that students who learn through the use of active learning techniques retain the material longer, achieve higher academic scores, and have lower failure rates than those who learn through more traditional lecture styles of teaching (Freeman et al., 2014). However, a limited understanding of active learning, lack of time, and large class sizes, among other challenges, can deter instructors from using active learning. To help promote statistics instructors' early, effective, and consistent use of active learning techniques, despite the challenges they face, it's important to understand the reasons why statistics instructors use active learning.

Drawing upon self-determination theory (SDT) and self-regulated learning (SRL) theory, this research provides a framework for understanding and investigating statistics instructors' reasons, or motivations, for using active learning. Through SDT, we use motivations and regulatory styles as means to understand how statistics instructors' use of active learning can be characterized along a continuum of self-determined behavior. Self-determined behavior, such as using active learning out of one's own interests or preference, is linked to increased curiosity, high degrees of effort, and continued long-term performance (Martens et al., 2004; Pinder, 2011), which may be beneficial in the promotion of and continued use of active learning in statistics classrooms. Consequently, helping foster the use of active learning as a more self-determined behavior may be the key to facilitating instructors' effective and consistent use of active learning in statistics classrooms.

We also discuss SDT's emphasis on how the fulfillment of autonomy, competence, and relatedness can make behavior, such as using active learning, more self-determined and, though the SRL cycle, can promote statistics instructors' learning about and initial use of active learning. Finally, we highlight the importance of and justification for using SDT to study statistics instructors' motivations to use active learning by comparing its key features to another commonly used motivation framework, Expectancy Value Theory.

Self-Determination Theory

Self-determination theory (SDT) (Figure 3.1) is a broad motivational framework that characterizes how self-determined a behavior is by describing why a person engages in a behavior, such as the use of active learning, through three distinct types of motivation: intrinsic motivation, extrinsic motivation, and amotivation. In the following sections, we discuss these three types of motivation, their associated regulatory styles, and how the fulfillment of autonomy, competence, and relatedness can promote motivations that characterize the use of active learning as more self-determined.

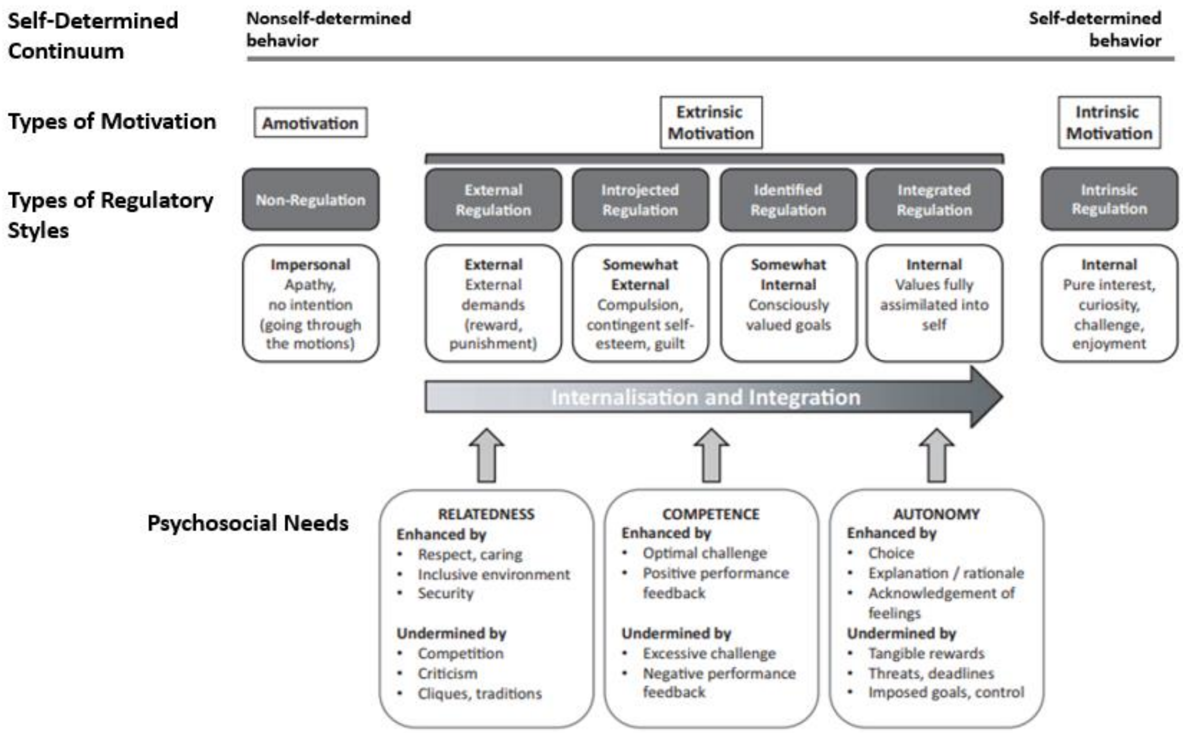


Figure 3.1: Self-determination theory framework. Adapted from "Motivation to learn: an overview of contemporary theories," by D.A. Cook and A.R. Artino, 2016, *Medical Education, 50*(10), p. 1010. Copyright Creative Commons Attribution License.

Types of Motivations

Intrinsic motivation characterizes fully self-determined behavior, where one engages in a specific behavior because they find the behavior enjoyable and satisfying, irrespective of any external factors. If a statistics instructor is intrinsically motivated to use active learning, they will show high levels persistence and effort when using such techniques (Martens et al., 2004).

On the opposite end of the spectrum, amotivation characterizes behavior that is least self-determined. This describes a state where one lacks any intention and purpose when engaging in a behavior—they just “go through the motions” (Ryan & Deci, 2000, pg. 72). If an instructor is amotivated to use active learning, it is likely they will not continue using it or will show minimal effort when doing so (Martens et al., 2004).

Extrinsic motivation characterizes behaviors as partially self-determined, where the reasons for engaging in a behavior are due to some external factor, such as rewards, punishments, guilt, or the value one awards to the outcome of the behavior. This type of motivation has been linked to a mix of behavioral outcomes that, depending on the external factors that drive the behavior, can either undermine or enhance more self-determined behavior (Ryan & Deci, 2000). Due to this complexity, Ryan and Deci (2000) categorize different types of extrinsic motivation through different regulatory styles.

Regulatory styles describe an individual’s emotions, values, beliefs, and attitude to help further explain why someone engages in behavior. Ryan and Deci (2000) identify six different regulatory styles: intrinsic regulation, integrated regulation, identified regulation, introjected regulation, external regulation, and non-regulation (Table 3.1). Intrinsic regulation and non-regulation uniquely describe intrinsic motivation and amotivation, respectively, and the other four regulatory styles help characterize extrinsically motivated uses of active learning as more or less self-determined behaviors (Figure 3.1).

Of the four regulatory styles that help characterize extrinsically motivated behavior, integrated and identified regulation are considered more autonomous forms of extrinsic motivation. An individual extrinsically motivated through an integrated regulatory style engages in active learning because the inherent value of the behavior aligns entirely with their own values. Extrinsically motivated behavior through an identified regulatory style suggests an individual uses active learning because they find the behavior important.

The remaining two regulatory styles, introjected and external, represent lesser self-determined behavior where an individual engages in behaviors purely based on external factors and influences. Introjected regulation characterizes an individual who imparts external consequences, such as feelings of shame or guilt, onto I to justify the use of active learning (Ryan & Deci, 2000). External regulation, however, characterizes the least self-determined type of extrinsic motivation, where an individual engages in active learning because they seek to avoid a punishment or achieve an award.

Type of Motivation	Regulatory Style	Classifications of self-determined behavior that are driven by emotions, values, beliefs, and attitudes.
Intrinsic Motivation	Intrinsic Regulatory Style	Interest, enjoyment, and satisfaction
Extrinsic Motivation	Integrated Regulation Style	Assimilation into personal values and beliefs
	Identified Regulation Style	Explicit value in behavior
	Introjected Regulation Style	Internal feelings of guilt or shame
	External Regulation Style	External pressures or rewards
Amotivation	Non-Regulation	“Going through the motions”

Table 3.1: The six regulatory styles that classify self-determined behavior.

All six of the regulatory styles help explain why statistics instructors use active learning. If the reasons an instructor uses active learning are described by identified regulation, integrated regulation, and/or intrinsic regulation, more positive, persistent behavioral outcomes are more likely to occur. The opposite is true for the other three regulatory styles that characterize behavior as less self-determined.

Changes in Regulatory Styles: Psychosocial Needs

Regulatory styles that describe the reasons why individuals engage in behavior are fluid, allowing the promotion of self-determined behavior. SDT identifies three psychosocial needs that need to be satisfied to promote behavior, such as using active learning, as more self-determined: autonomy, relatedness, and competence (Figure 3.1). Autonomy refers to the ability to make free choices and be in self-control (Keller, 2016). One's competence refers to having sufficient knowledge or skill to engage in an activity (Von Treuer & Reynolds, 2017). Relatedness is the desire to be connected to others (Ryan, 1993). To foster active learning as a more self-determined behavior, SDT suggests providing statistics instructors the opportunity to perceive they are freely making the choice to use active learning, increasing their perceived competence in using active learning in the classroom, and developing a space where they can relate with others about using such techniques.

When an instructor's need for autonomy is filled, the instructor is more likely to provide students with meaningful tasks, share the responsibility of learning with students, and create opportunities for students to make decisions (Reeve, 2002; Ryan & Deci, 2000). Similarly, when instructors' needs for competence are met, they are more apt to engage with, use, and persist with student-centered forms of teaching (e.g., active learning) (Brenner, 2022; Schellenbach-Zell & Gräsel, 2011). As instructors gain a heightened sense of relatedness with both colleagues and their students, they are more likely to report higher levels of teaching engagement, including the construction of more instructor-student exchanges (Evelein, Korthagen, & Brekelmans, 2008).

The fulfillment of these psychosocial needs manifests more self-determined behavior through the mental processes of internalization and integration (Figure 3.1) (Ryan & Deci, 2000). Internalization and integration are mental processes where the individual identifies with the values of engaging in a specific behavior and assimilates this into their self and personality (Ryan & Deci, 2000). For example, an instructor's initial reasons for using active learning may be described by external regulation before they later identify and accept the value of active learning, changing to a regulatory style that characterizes more self-determined use of active learning.

Self-Regulated Learning

SDT is a broad motivational framework that can help us better understand how the use of active learning can be characterized as more or less self-determined, and outlines how using active learning can be fostered as a more self-determined behavior. Self-regulated learning (SRL) theory, through its relationship with SDT, offers a model for facilitating instructional change, such as including the use of active learning techniques in the classroom. This model may be particularly useful for studying and promoting the use of active learning for newer statistics instructors who may be learning about active learning for the first time.

As illustrated in Figure 3.2, SRL theory hypothesizes a cycle comprised of three phases (plan/forethought, learning/performance, and evaluate/self-reflection) in which statistics instructors can transform information they are learning about active learning into heightened mental abilities (Zimmerman, 2000). For example, during the forethought phase, statistics instructors could analyze information about active learning that they already know, set goals about using such techniques, and plan out how to achieve those goals. During the performance phase, instructors would then use, or attempt to use active learning in their classroom, monitoring both their use of and the students' reactions to such techniques by recording observations or journaling notes. Finally, during the self-reflection phase, the statistics instructors would reflect on these experiences and observations, assessing how they performed in their use of active learning and identifying reasons for each success or failure. Engaging in this SRL cycle leads to self-awareness, self-assessment, and openness to change (Zimmerman, 2000), so learning about active learning in this way may be a key component to fostering early use of active learning for statistics instructors who may otherwise teach using more traditional types of methods, such as lecture.

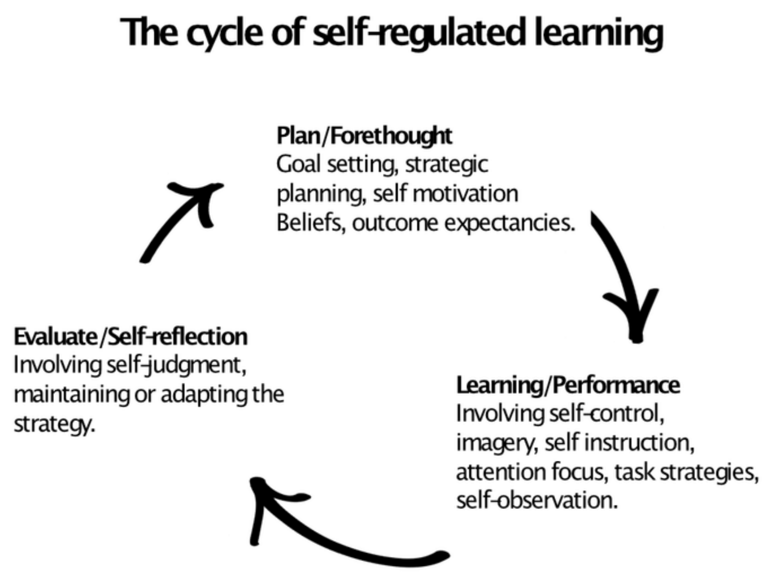


Figure 3.2: The self-regulation learning cycle that describes a learning model through three phases. Adopted from "Why Medical Students Choose to Use or Not to Use a Web-Based Electrocardiogram Learning Resource: Mixed Methods Study," by N. Nilsson et al, 2019, p. 9. Copyright Creative Commons Attribution License.

Research suggests that instructors may be more likely to embrace self-regulated learning if their psychosocial needs, outlined in SDT, are fulfilled. Similar to how the psychosocial needs identified in SDT help support the internalization and integration of active learning to characterize such use as more self-determined, these needs, when supported through the SRL cycle, help foster changes in instructors' use of active learning. In practice, this model acts as a framework for how to teach active learning to statistics instructors through the SRL cycle and emphasizes the need to foster an environment that satisfies instructors' needs for autonomy, competence, and relatedness to help facilitate the use of active learning techniques over instructor-centered forms of teaching.

Other Motivational Frameworks: A Comparison

We view statistics instructors' use of active learning through SDT, where the reasons why instructors use active learning are classified by different motivations and regulatory styles. SDT uses these motivations and regulatory styles to characterize how self-determined a behavior, such as an instructor's use of active learning, is on a continuum. When a behavior, such as the use of active learning, is more self-determined, a statistics instructor will tend to find active learning interesting, persist in using active learning, and engage more with active learning as a result (Brenner, 2022).

For an instructor's use of active learning to become more self-determined, SDT theorizes that their needs for competence, relatedness, and autonomy must be fulfilled. These needs, in conjunction with the SRL cycle, provide a model to foster change in instructional behavior geared more towards the use of active learning.

Other motivational theories, such as expectancy value theory (Wigfield & Eccles, 2000), exist for framing statistics instructors' motivations to use active learning. Expectancy value theory (EVT) suggests that motivations to use active learning are explained by a statistics instructor's perceived value towards active learning, their expectations to succeed in using active learning, and the opportunity cost associated with using active learning in their classroom. Expectancy relates to when an instructor perceives they will be good at using active learning. Value is broken down into intrinsic value, i.e., how much you enjoy the task; utility value, i.e., how much you believe the task is worth it; and attainment value, i.e., how much you believe this task is important. Cost is measured by how an instructor may perceive active learning to be "worth it" compared to other instructional activities.

Although there are similarities between SDT and EVT, they differ in their characterizations of how motivations change. SDT theorizes that the degree to which one's needs of autonomy, competency, and relatedness are fulfilled impacts their motivations to engage in a behavior, whereas EVT emphasizes how much someone values a behavior, expects to succeed at it, and how much the behavior will cost (i.e., opportunity cost). Existing literature suggests that instructional practices, such as the use of active learning, can change as instructors feel more autonomous and connected with others (Herman, West, & Hahn, 2015; Lyle & Peurach, 2022). Unlike EVT, SDT explicitly acknowledges these needs of autonomy and relatedness, providing a framework for understanding how they impact statistics instructors' motivations to use active learning. Measuring, modeling and understanding this type of impact is particularly important given the fluid nature of one's motivations and instructional practices over time (Ostinelli, 2016). Uncovering how statistics instructors' needs of autonomy and relatedness can be fulfilled may offer professional development programs additional insight into how to promote and support statistics' instructors early, effective, and consistent use of active learning through the fulfillment of such needs.

Framing statistics instructors' motivations to use active learning through SDT instead of EVT also supports the broad study of both experienced and newer statistics instructors. Unlike SDT, EVT measures achievement motivation which theorizes that an instructor uses active learning because they desire to excel or master this more advanced teaching technique. However, some statistics instructors, such as graduate student instructors (GSIs), teach for a finite amount of time and do not pursue a teaching career after graduation. Given this short window of time, it may not make sense to study GSIs' use of active learning through achievement motivation. If a GSI does not plan to pursue a teaching career, they may not approach the use of active learning through the lens of achievement. Additionally, a GSI often has little to no experience with teaching or using active learning. Forming an expectation around how successful one is when using active learning takes time; expectations come from experiences that not all statistics instructors may have. By using self-determined behavior and psychosocial needs instead of achievement motivation to frame statistics instructors' motivations to use active learning, SDT applies to both experienced and newer statistics instructors, not just those who have already formed expectations of success. Additionally, SDT's natural extension to the SRL cycle lends itself to investigations into how to foster statistics instructors' initial use of active learning, offering insight that may be useful for supporting newer statistics instructors' professional development.

Discussion and Conclusions

Statistics instructors are encouraged to use active learning techniques in their classrooms. However, national recommendations to use these techniques do not guarantee their use, let alone their effective use. We propose the use of SDT and SRL theory as a framework to better understand how to facilitate early, effective, and consistent use of active learning techniques in statistics classrooms.

The SRL cycle, in conjunction with autonomy, competence, and relatedness, provides a framework for fostering statistics instructors' initial use of active learning. We encourage the use of this framework to inform future recommendations on how to support statistics instructors' early use of active learning techniques. Similarly, we propose the use of SDT as a tool to inform how to support effective and consistent uses of active learning when teaching statistics. SDT details the sense of self-determination a statistics instructor feels when using active learning, highlighting important predictive behavioral outcomes, including their effort, effectiveness, and persistence with such techniques. Motivations described by SDT (intrinsic motivation, extrinsic motivation, and amotivation) characterize why statistics instructors are using active learning along a continuum of self-determined behavior. Understanding how self-determined the use of active learning is for statistics instructors may inform future recommendations on how to foster consistent and continued use of active learning.

MOTIVATION MATTERS: DEVELOPING AN INSTRUMENT TO MEASURE NEWER INSTRUCTORS' SITUATIONAL MOTIVATION TO USE ACTIVE LEARNING IN THE CLASSROOM WHEN TEACHING STATISTICS

Contribution of Authors and Co-Authors

Author: Elijah S. Meyer

Contributions: Designed study, collected data, performed analyses, interpreted results, and wrote manuscript.

Co-Author: Jennifer L. Green

Contributions: Provided guidance on all aspects of the study and edited drafts of manuscript.

Co-Author: Stacey A. Hancock

Contributions: Discussed framework and edited drafts of manuscript.

Manuscript Information

Elijah S. Meyer & Jennifer L. Green & Stacey A. Hancock

Statistics Education Research Journal

Status of Manuscript:
X___ Prepared for submission to a peer-reviewed journal
____ Officially submitted to a peer-reviewed journal
____ Accepted by a peer-reviewed journal
____ Published in a peer-reviewed journal

Publisher: *International Association for Statistical Education & International Statistical Institute*

Abstract

National recommendations for teaching statistics include the use of more advanced teaching techniques, such as active learning, to better engage students in the learning process. However, the adoption of active learning techniques remains limited, especially for newer statistics instructors. Understanding how self-determined their use of active learning is through their motivations to use active learning can help inform how to support newer instructors' adoption of such techniques. Thus, we designed a five-phase study to develop an instrument that measures newer statistics instructors' types of motivations to use active learning. During the first three phases, we created the instrument items and obtained evidence of content validity. We then used data from newer statistics instructors in Phases 4 and 5 to obtain evidence of the instrument's construct validity and reliability. The resulting reliability and validity evidence suggest that this instrument may help support future studies' investigations of motivation, providing researchers and statistics teacher educators a new tool for measuring and understanding why newer statistics instructors use active learning.

Introduction

Statistics is widely recognized as one of the most important quantitative subjects available through university curricula (Weihs & Ickstadt, 2018). Statistics curricula allow students to learn appropriate applications of statistical methods that can empower individuals to make discoveries and decisions with data. To ensure students' understanding of these applications, there has been a call to transform instructional practices from more traditional, instructor-centered styles of lecture to more engaging, student-centered practices, such as the use of active learning techniques. Active learning techniques are any classroom-based practices that engage students in their learning through means of problem solving or discussions (Nguyen et al., 2021). Examples of active learning include large-group discussions, think-pair-share, and the use of hands-on technology (Center for Research on Learning and Teaching, n.d.). Students taking science, technology, engineering, and mathematics (STEM) courses tend to earn higher exam scores and have lower failure rates when they learn through the use of active learning techniques compared to more traditional lecture styles of teaching (Freeman et al., 2014). Yet, the adoption of active learning techniques in statistics classrooms remains inconsistent (Apkarian et al., 2021). This is especially true when the class is taught by a newer statistics instructor, such as a graduate student instructor (GSI) (Justice, 2020).

Instructors' motivations to teach are related to which types of teaching practices they use (Han & Yin, 2016). However, there is no current way to measure newer statistics instructors' motivations to use active learning. Given the documented benefits of using active learning techniques (Freeman et al., 2014) and the critical roles newer statistics instructors such as GSIs serve in teaching introductory statistics courses (Friedman, 2017), it is important to be able to measure the degree to which newer statistics instructors are motivated to teach using such techniques. Professional development programs often are created to support newer statistics instructors' use of more advanced teaching techniques, such as active learning. Understanding why newer statistics instructors use active learning, could inform how these programs better support instructors' adoption of such techniques.

Motivational Constructs

Motivations we intend to measure are understood through the lens of self-determination theory (SDT). Proposed by Deci and Ryan (1985; 2000), this framework outlines why individuals engage in behaviors through three types of motivation: intrinsic motivation, extrinsic motivation, and amotivation (Figure 4.1). A newer statistics instructor who is intrinsically motivated to use active learning does so because they find it fun or enjoyable; their reasons for using active learning are inherently satisfying and are not driven by any external factors. In contrast, instructors who are extrinsically motivated to use active learning are driven by external factors, including wanting to avoid punishment or earn rewards, or seeing value in the use of active learning. Amotivation is the absence of both intrinsic and extrinsic motivation, where someone is disinterested and just "going through the motions."

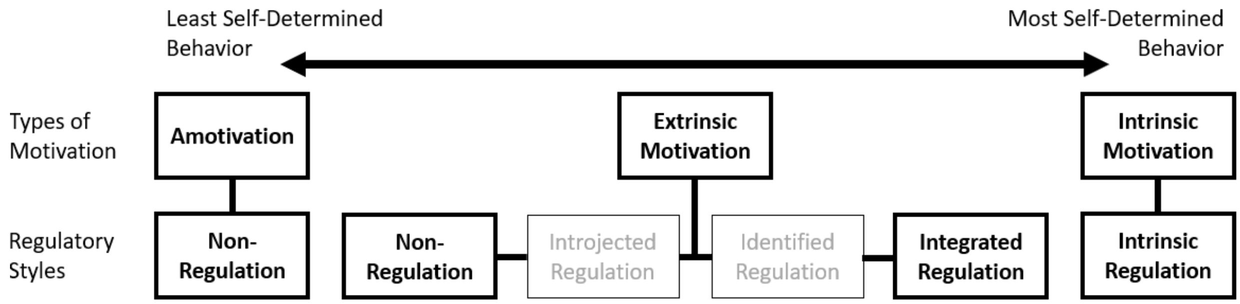


Figure 4.1: The three types of motivation and six regulatory styles outlined by self-determination theory. The types of motivation and regulatory styles selected to be included on our instrument are bolded.

Because reasons for engaging in behavior are complex, intrinsic motivation, extrinsic motivation, and amotivation are further divided into six subsequent regulatory styles (Table 4.1). These six regulatory styles refer to how an individual's emotions, values, beliefs, and attitudes shape their reasons for engaging in a behavior, such as using active learning. Intrinsic regulation, which is uniquely classified by intrinsic motivation, characterizes when a newer statistics instructor's use of active learning is driven by personal satisfaction. Extrinsic motivation is classified into four different types of regulatory styles. A newer statistics instructor who is extrinsically motivated to use active learning through an integrated regulatory style does so because the use of active learning aligns completely with their values as an instructor. If classified through an identified regulatory style, the instructor is using active learning because they find the use of such techniques important, whereas behavior that is extrinsically motivated by guilt and shame characterizes an introjected regulatory style. External regulation describes a statistics instructor who uses active learning because they feel pressure from others or perceive that they will gain a reward if they use such techniques. Non-regulation is uniquely classified by amotivation and suggests that a newer statistics instructor lacks intentionality and purpose when using active learning.

Regulatory Style	**Description**
Intrinsic Regulation	Interest, enjoyment, and satisfaction
Integrated Regulation	Assimilation into personal values and beliefs
Identified Regulation	Explicit value in behavior
Introjected Regulation	Internal feelings of guilt or shame
External Regulation	External pressures or rewards
Non-Regulation	"Going through the motions"

Table 4.1: The six regulatory styles and the corresponding emotions, values, beliefs, and attitudes characterizing each.

These six regulatory styles categorize behavior along a self-determination continuum (Figure 4.1). Self-determined behavior is behavior in which a person perceives to be in control, to make their own choices, and to manage their own lives. When the use of active learning is more self-determined, a statistics instructor will tend to find active learning interesting and show increased persistence when using active learning in their classroom (Ryan & Deci, 2000; Vallerand, 1997).

In this research, we aim to measure four of the six regulatory styles: intrinsic regulation, integrated regulation, external regulation, and non-regulation (Figure 4.1). By measuring these different regulatory styles, we hope to differentiate between different levels of newer statistics instructors' self-determined behavior and explore more nuanced reasons for why they use active learning than measuring only the three types of motivation. We chose the regulatory styles that classify the most and least self-determined behavior—intrinsic regulation and non-regulation—due to their stark differences in potential behavioral outcomes. We reference each by their respective type of motivation (intrinsic motivation and amotivation, respectively). Within extrinsic motivation, which is classified by four different types of regulatory styles, we again chose the regulatory styles that classify the least (external regulation) and most (integrated regulation) extrinsically motivated self-determined behavior and reference each by their respective names throughout the manuscript.

Hierarchical Levels of Motivation

Motivation is understood hierarchically through three levels: global, contextual, and situational (Figure 4.2). The types of regulatory styles we intend to measure can be measured at each of these three levels. The highest level in the hierarchy is the global level, which considers how environmental factors influence an individual's intrinsic motivation, extrinsic motivation and amotivation (Núñez & León, 2019). Measuring motivation at the global level assesses one's overall orientation towards a behavior (e.g., "In general, I do things because..."). The middle, contextual level models how related activities or social factors, such as one's education level, leisure time, and state of their interpersonal relations, may influence their motivation to engage in behaviors (Blais, Vallerand, Brière, Gagnon, & Pelletier, 1990). The bottom, situational level of the hierarchy addresses situational motivation, or the motivation individuals experience when engaging in an activity. This type of motivation refers to the "here and now," explaining why an individual is currently engaging in an activity (Vallerand, 1997).

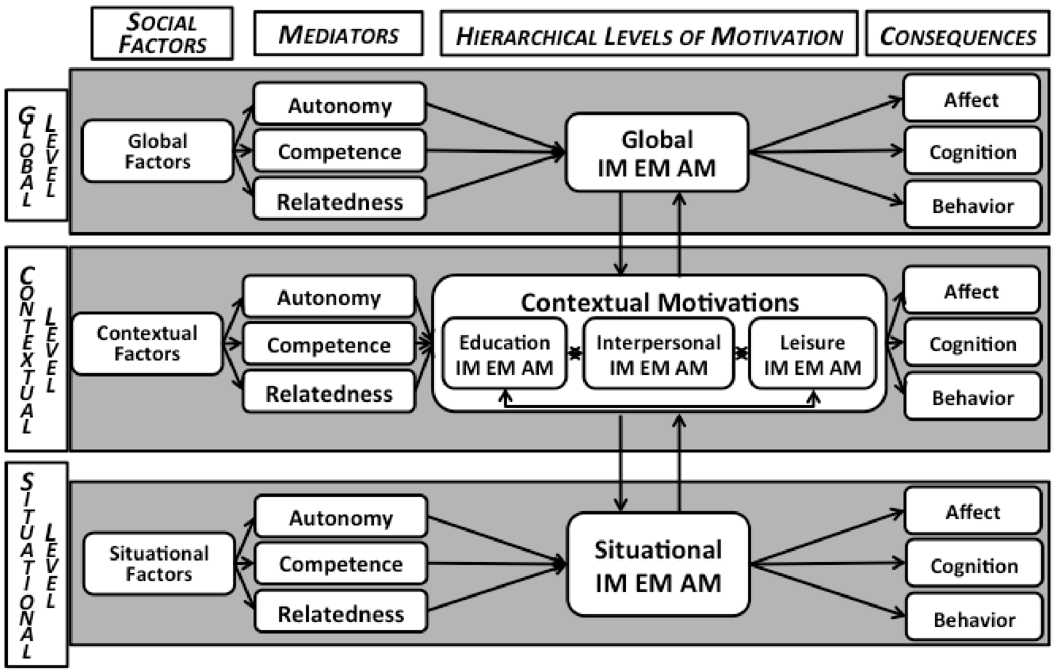


Figure 4.2: The Hierarchical Model for Motivation that explains motivation at three levels: global, contextual, and situational. Adopted from "Toward a hierarchical model of intrinsic and extrinsic motivation," by R. J. Vallerand, 1997, In M. P. Zanna (Ed.), Advances in experimental social psychology, 27, p. 271-360. Copyright Creative Commons Attribution License.

In this research, we chose to measure the four regulatory styles identified earlier at the situational level. We argue that measuring motivation and classifying behavior along a self-determined continuum is critical when studying newer statistics instructors' motivation to use active learning. Studying motivation at specific time points during "the here and now" provides more personalized information about newer statistics instructors' use of active learning than studying it at higher contextual and global levels. In addition to offering a deeper understanding of statistics instructors' motivations, this detailed information may also create a foundation to help investigate how professional development decisions deter or help foster the use of active learning as more or less self-determined. This information can inform practice, such as the development and refinement of professional development to help support more self-determined uses of active learning.

Research Aim

Although motivational instruments already exist, none are suitable for measuring newer statistics instructors' motivation to use active learning. The Work Tasks Motivation Scale for Teachers (Fernet, Senécal, Guay, Marsh, & Dowson, 2008) and the Teacher Motivational Assessment Scale (Obunadike, 2013) both measure instructors' motivations when teaching, but neither instrument focuses on the use of active learning or is designed for use with newer statistics instructors. Other instruments, such as the Intrinsic Motivation Inventory (McAuley, Duncan, & Tammen, 1989) and the Situational Motivational Scale (Guay, Vallerand, & Blanchard, 2000), measure types of motivations more generally for individuals in field and laboratory settings. However, again, neither instrument has evidence of validity or reliability for use with newer statistics instructors. Further, none of the listed instruments simultaneously measure all the constructs of interest for this study. In this research, we aim to develop an instrument that measures newer statistics instructors' intrinsic motivation, integrated regulation, external regulation, and amotivation to use active learning. Understanding why newer statistics instructors use active learning, who are often the population of professional development, could inform how to better support their adoption of such techniques. We define newer statistics instructors as any statistics instructor with three or less years of teaching experience, as well as any graduate student instructor. To measure motivation, newer statistics instructors must be using active learning when teaching statistics.

Research Design

To meaningfully measure our four motivational constructs (intrinsic motivation, integrated regulation, external regulation, and amotivation), we developed a five-phase research plan to construct our instrument and establish initial evidence of content validity, construct validity, and reliability (Table 4.2). Content validity refers to the extent to which our items are representative of the types of motivations we intend to measure, in their entirety. Construct validity is the degree to which our instrument can measure each motivational construct as intended and defined by SDT. Reliability pertains to the extent sets of items measure the same intended motivational construct.

	Phases and Procedures	Evidence
Instrument Creation	Phase 1: Item Development • Instrument Blueprint • Expert Opinion • Review of Existing Instruments	Content Validity
	Phase 2: Expert Item Review and Scale Development • Expert Opinion • Scale Development and Review	Content Validity
	Phase 3: Cognitive Interviews • Semi-Structured Interviews	Content Validity
Pilot Studies	Phase 4: Pilot 1 • Summary Statistics • Graphical Assessments • Cross-Correlations • Cronbach's Alpha	Construct Validity Reliability
	Phase 5: Pilot 2 • Proportion of Non-Selected Response Options • Correlation Matrix • Cross-Correlations • Confirmatory Factor Analysis • Coefficient Omega	Construct Validity Reliability

Table 4.2: Five-phase research design for instrument creation and pilot studies.

Instrument Creation

Phase 1: Item Development

In Phase 1, we developed an instrument blueprint to guide our item writing process and reviewed existing instruments to establish initial evidence of content validity, i.e., that our items assess all aspects of each motivational construct we intended to measure (Menold, Jablokow, Purzer, Ferguson, & Ohland, 2015). We also solicited expert feedback about different active learning techniques and used that information to choose one active learning technique on which to measure motivations. The following sections provide more information about each of these steps and how they informed initial item development.

Item Blueprint Using definitions from SDT (Ryan & Deci, 2000), we created an item blueprint that provided working definitions of each type of motivational construct we intended to measure (Appendix E). The blueprint also included working definitions of different active learning techniques. Because different active learning techniques may elicit different types of motivations and may be interpreted differently across newer statistics instructors, we selected and defined four active learning techniques as potential options on which to measure motivations. Based on our review of relevant literature, statements, and recommendations on how to use active learning techniques in the classroom (e.g., Carver et al., 2016), all authors came to a consensus on which four active learning techniques to consider—group work, using technology, large group discussions, and using real data—and agreed upon their definitions.

Expert Opinion To select which of the four active learning techniques to focus our instrument on, we sought expert opinion. Someone was classified as an expert if they had extensive research on or experience teaching statistics, or if they had experience working with newer statistics instructors. We surveyed experts on the types of active learning techniques they would include on an instrument that measures motivations. The survey contained the four different active learning techniques and their definitions from our item blueprint (Appendix F).

Of the 30 experts we contacted, 13 across 12 different institutions of higher education agreed to participate and completed the survey. These institutions of higher education included both public and private institutions, as well as Research 1 universities. Eight institutions reside in the Midwest, while the others are in the Northwest, Central, and Pacific regions of the United States.

The experts had approximately four weeks to complete and return the survey over email. When reviewing the list of active learning techniques provided, experts had the option to add any additional active learning techniques and provide working definitions of them. Some experts listed techniques such as reflective writing and classroom polling. After the experts listed additional active learning techniques, each was asked to critique the working definitions of all the techniques and rank all techniques (researcher provided and expert created) in order from most to least important for an instrument measuring newer statistics instructors' motivations.

For each active learning technique, we summarized experts' feedback and identified common comments and suggestions. For example, experts agreed that there needed to be more specifics on what constitutes technology and what does not. We then used this feedback to update our definitions of the active learning techniques in the item blueprint. In addition, we summarized the final rankings of the four active learning techniques we provided on the survey (Table 4.3), along with the rankings of the nine additional active learning techniques that experts provided; few of the additional active learning techniques that experts provided were ranked higher than third priority, so they are not included in Table 4.3. Note that not all the columns sum to 13 because some experts did not rank each of the four techniques provided. Based on these rankings and experts' feedback, we chose to focus our instrument and write items in the context of newer statistics instructors' motivations to use group work and technology.

Priority	Group Work	Technology	Real Data	Large Group Discussions
1	6	0	1	1
2	3	6	2	0
3	0	2	3	2
4	2	3	1	6
5	0	0	5	1

Table 4.3: Frequency of how experts prioritized the listed active learning techniques.

Review of Existing Instruments When writing items, we continuously referenced the definitions on the item blueprint to ensure our items represented characteristics of their respective motivational construct and active learning technique. We also reviewed multiple existing instruments to identify additional aspects of each motivational construct we intended to measure and highlight language that is important to the context of Statistics (Appendix E). Motivation instruments were predominantly from the field of Psychology and had items relating to at least some of the four motivational constructs we identified. We also studied Statistics instruments to better incorporate the context and language of Statistics into our instrument. Using this information, along with item writing guidelines recommended by Haladyna and Rodriguez (2013), we wrote 34 items in total across our four constructs in the context of group work and 34 in the context of using technology.

To prepare the initial instrument for review, we grouped items for each of the contexts by construct to help mitigate the cognitive load of those willing to review and take the instrument in future phases. We also created background questions, placing questions used to identify those in the intended population of interest at the beginning of the instrument (e.g., Are you a graduate student? Are you teaching statistics?) and other questions about newer statistics instructors' characteristics (e.g., prior experiences teaching, prior training programs, classroom context) at the end of the instrument.

Phase 2: Expert Item Review and Scale Development

In Phase 2, we established further evidence of content validity for our items through two rounds of expert review. These initial reviews of item interpretability helped determine whether the items were understood as intended. We also adopted the response scale from the Situational Motivational Scale (Guay et al., 2000) and asked experts questions about the scale's appropriateness for our context. We then triangulated existing literature with expert responses to create our own response scale for the instrument.

Expert Item Review

To solicit experts' feedback on the interpretability of our items, we conducted two rounds of expert item review. In Round 1, we created and sent out a survey (Appendix G). Five experts from five different institutions, separate from those in Phase 1, were contacted and asked to complete the survey. Each expert was informed that all information would remain anonymous to all but the authors of this paper and that they were not obligated to answer any of the survey questions. Each expert had approximately two weeks to respond and return the survey over email. All five experts agreed to participate and completed the survey.

The survey contained working definitions of the active learning strategies selected, working definitions of the motivational constructs to be measured, and a draft of our items. Each expert was asked to provide feedback on three specific areas: the active learning strategies' definitions, the drafted items, and the motivational constructs' relationship with the drafted items. These questions came after each set of items for a motivational construct. Example questions included: Do you believe each item reflects qualities of intrinsic motivation; Do you believe there are missing items; Are aspects of intrinsic motivation's definition mis- or underrepresented within the set of items; Are there concerns about the wording of the item (e.g., double-barreled, idioms, jargon, etc.)? Additionally, spaces were provided for experts to add any comments at the end of each section.

We conducted a similar expert item review process for Round 2. Six experts from Phase 1 were recruited and completed the same survey we had used in Round 1. Each expert had the option to provide alternative oral feedback through a video call. No expert chose this option. Each expert was randomly assigned to review either group work or use of technology items.

For each item, we summarized experts' feedback and identified common comments and suggestions. The experts' feedback helped us identify opportunities to improve item clarity and content validity so that the items better aligned with each construct's working definition. Feedback and changes to the instrument items included: refraining from wording items to have a "right answer," discerning verbiage more associated with beliefs than motivations, and incorporating consistent phrasing.

Scale Development During the second round of item review, we also asked experts questions about the seven-point Likert scale included with the items. We adopted the scale from the Situational Motivational Scale (Guay et al., 2010) (7- Corresponds exactly, 6- Corresponds a lot, 5- Corresponds enough, 4- Corresponds moderately, 3- Corresponds a little, 2- Corresponds a very little, 1- Corresponds not at all). After reviewing the experts' feedback and literature, we chose to instead use a six-point Likert scale with the following response options: Strongly agree, Agree, Slightly agree, Slightly disagree, Disagree, and Strongly disagree. To arrive at this scale, we contemplated the use of a midpoint and the use of a different number and verbiage of response options. We justify the choice of response option verbiage, and the decisions to remove the midpoint and use a six-point scale below.

When answering questions about the initial seven-point Likert scale, experts raised concerns that the response option verbiage was extremely subjective and may be interpreted differently across individuals. For example, one expert remarked, "I'd reconsider the words, 'a little', 'moderately', 'enough' because those are very subjective [when answering these items]. What I consider 'enough' could be very different from what another considers as 'enough.'" Hancock and Volente (2020) emphasize the importance of language, suggesting that even when referring to common terminology, individuals do not always mean or may not interpret the exact same thing as others. Lee et al. (2002) claim that more subjective language may also be interpreted differently across cultures. Being mindful of the language used in the scale response options may lead to an overall better measure of the intended constructs.

We also chose to not include a midpoint response, as the misuse of the midpoint compromises the data collected and may misrepresent the construct intended to be measured. Research shows that individuals do not always interpret a neutral option as a way to express a nonaligned answer between agreement and disagreement, and they may select a midpoint even if this does not reflect their true stance towards a specific item (Chyung, Roberts, Swanson, & Hankinson, 2017). Additionally, individuals may use the midpoint as a “dumping ground” when responding to items (Kulas, Stachowski, & Haynes, 2008). This may occur in a variety of situations, including when respondents are unfamiliar with items, when items are ambiguous, or when survey fatigue sets in. In a study conducted by Nadler et al. (2015), 635 college students described how they would interpret the midpoint “Neither agree nor disagree” on a five-point Likert scale, and many responded with the following interpretations: “Don’t care” (14%) and “Unsure” (13%), which are misuses of a midpoint.

Although other research suggests one can avoid the misuse of a midpoint by improving item wording or presenting alternative options such as “I don’t know” to catch “dumping ground” answers (Kulas & Stachowski, 2013), Garland (1991) found that individuals were more likely to select the midpoint instead of a more negative option if the context was socially undesirable. For our research, it is reasonable to assume that newer statistics instructors may be more likely to select a midpoint than a negative option, even if the midpoint does not reflect their true motivation to use active learning, due to an idea that it is socially undesirable to not follow current teaching recommendations. In choosing to remove the midpoint, we remove the issue of midpoint misinterpretation. The exclusion of the midpoint forces individuals to think critically about the options they select in a forced choice format which may provide a more authentic data collection process.

We chose six response options to ensure respondents had the opportunity to answer an item in a way that most aligned with how they felt. Preston and Colman (2000) revealed that scales with lower numbers of response options (e.g., 2, 3 and 4) reported lower scores to the question, "Were you allowed to express your feelings adequately" than scales with six options. Further, when comparing four- and six-point scales, Chang (1994) found no effect on criterion-related validity and Leung (2011) found the instruments had similar psychometric properties.

Phase 3: Cognitive Interviews

During Phase 3, we conducted cognitive interviews with six newer statistics instructors to continue gathering evidence of content validity of our items. These interviews helped us explore how newer statistics instructors interpret and respond to the items and, before being conducted, involved three different steps of preparation: developing an interview protocol, recruitment, and interviewer training (Peterson, Peterson, & Powell, 2017).

Interview Protocol We developed a semi-structured interview protocol with a series of different probes (questions) to help facilitate conversation with the interviewee and collect rich information about the interpretability of our items (Appendix H). We included three different kinds of probes: expansive, anticipated, and conditional. Expansive probes are often used to investigate respondents' mental processes and may include questions such as "Tell me more about that" or "How did you come up with that?" By asking expansive probes, we could then construct specific follow-up probes about what a respondent said, allowing us to further investigate the respondent's mental process for interpreting a term or item. Because specific follow-up probes are not typically included on an interview protocol, the first author constructed them in the moment based on what a respondent said during their interview.

When creating the interview protocol, we also included anticipated and conditional probes. Anticipated probes are questions about topics or phrases with which the interviewer anticipates a respondent will have difficulty interpreting, and conditional probes are questions that are triggered by verbal or nonverbal cues from a respondent. For example, the interviewer may ask, "Why did you answer this item this way," if the respondent chooses an extreme response. On the interview protocol, we included probes about item words or phrases that we anticipated would be difficult to interpret. We also crafted conditional probes to investigate inconsistent item responses, unique statements, or nonverbal cues newer statistics instructors made when answering items.

When writing the probes, we designed them to not have a correct answer so that respondents would not be influenced to answer in a certain way. We also created a flexible interview protocol so that we could ask unscripted, emergent probes during the interview process. These unscripted questions enabled the first author to ask about unique information a respondent shared, such as why a question took a long time to answer or was answered differently than other similar items, and to obtain potentially valuable information for better understanding the interpretability of an item.

Interviewer Training Before conducting the cognitive interviews, the first author held a mock interview with a graduate student volunteer, paying specific attention to the graduate student's verbal and nonverbal communication. Specifically, the first author paid attention to and recorded when the graduate student looked confused, contradicted themselves, or was reluctant to answer certain items (Beatty & Willis, 2009). These observations were listed on the interview protocol to help the first author identify important verbal and nonverbal cues when conducting the cognitive interviews. When observed, these cues would offer initial evidence that an item was difficult to interpret or unclear.

Recruitment We used criterion sampling to purposefully select a diverse sample of newer statistics instructors who were using active learning to teach a statistics class. Because we wanted to ensure multiple perspectives and experiences were represented during data collection (Ames et al., 2019), we selected participants from the intended population based on the following sampling criteria: (1) beginning newer statistics instructor (≤ 3 terms of teaching); (2) experienced newer statistics instructor (> 3 terms of teaching); (3) international newer statistics instructor; and (4) non-international newer statistics instructor. These criteria were purposefully selected based on literature that novice and more experienced instructors may be in different stages of teaching and may experience active learning differently (Auerbach et al., 2018). Further, we prioritized having a more diverse perspective on item interpretation by recruiting both international and non-international newer statistics instructors. These diverse perspectives allowed us to investigate whether newer statistics instructors were interpreting the items both consistently and as intended. They also allowed us to make our items more relatable and relevant across a spectrum of newer statistics instructors.

To recruit newer statistics instructors for the cognitive interviews, we emailed faculty who supervise newer instructors that teach a statistics course. The email included a recruitment email template that faculty were invited to edit and send to their respective newer instructors for considering participation in our research study. The first author also directly contacted newer statistics instructors using the recruitment email. The recruitment email asked those who were interested in participating to answer questions about whether they were currently teaching a statistics course and using active learning strategies, as well as questions related to the four sampling criteria. From those who were within our intended population and fit any of the specified sampling criteria, we planned to select newer statistics instructors at different locations to help maximize the diverse perspectives of those recruited.

We selected six newer statistics instructors from three different universities to participate in cognitive interviews. These universities were located within the Northwest, Midwest, and Northeast regions of the United States. All the participants were graduate students at the time of recruitment, with only one of them, an international student, having two or less terms of teaching experience (Table 4.4).

	International newer statistics instructor	**Non-International newer statistics instructor**
Beginning newer statistics instructor	1	0
Experienced newer statistics instructor	2	3

Table 4.4: Breakdown of sampling criteria across six participants for cognitive interviews.

After selecting the participants, we emailed them a video consent form and asked to schedule a video call for the cognitive interview. We also emailed each participant a Qualtrics link to the instrument items before their interview started.

Conducting Interviews Cognitive interviews were conducted and recorded with permission using Zoom to cross check notes taken during the interview. During the cognitive interviews, the first author requested that participants read each item aloud and articulate why they selected their particular response. The first author asked emergent questions whenever participants had difficulties answering a question or articulating their reasoning. At the end of each page of items, participants were asked questions from the interview protocol to further discuss item interpretability.

During the interview, the first author wrote memos of important ideas discussed and any verbal or nonverbal cues, such as wait time, expressed by the newer statistics instructors when responding to the items. This information was used to flag items that may have been hard to answer, hard to interpret, or irrelevant for newer statistics instructors. Flagged items were then grouped together across all respondents to assess if any items were flagged by more than one respondent. We then summarized respondents' comments and feedback to identify any emerging patterns within items.

After the cognitive interviews, multiple authors discussed how the participants interpreted the items and considered evidence for how the items should be updated. Participants' responses from the cognitive interviews suggested that it was not appropriate to have items that related to group work and technology on the same instrument, as all six participants answered one set of items relative to the other. After noticing this, the lead author asked each participant emergent questions to help determine which active learning technique would be most appropriate to include on the instrument. Participants suggested that group work may be the most appropriate, because the use of technology can have variable interpretations and may not be as accessible to newer statistics instructors as using group work. Additionally, experts from Phase 1 voted group work as the active learning technique they would most prioritize on an instrument that measures motivations. Based on this information, we chose to focus our instrument on group work.

Participants' responses from the cognitive interviews also helped us improve the items' interpretability. During the cognitive interviews, multiple participants answered items with the disclaimer, "If active learning worked." These items were labeled as "outcome dependent," because they relied on the outcome that active learning worked. Some examples of outcome-dependent items included, "I feel pleased about lessons where I have used group work" and "I get excited when using group work to teach statistics." To mitigate this dependency, we introduced phrases such as "typically" at the beginning of such items. Additionally, we removed items that were consistently hard to interpret or misinterpreted, such as “Using group work to teach statistics aligns with my identity as an instructor.” We also grouped together items that were interpreted in the exact same way. For example, many individuals interpreted items with the terms “enjoyable” and “interesting” similarly, interpreting both terms as “fun.”

After Phases 1 through 3, a total of 34 items were created, selected, and revised to be included in Phase 4, the first pilot phase of our research study. The initial pilot instrument had 8 intrinsic motivation items, 11 integrated regulation items, 7 external regulation items, and 8 amotivation items. We then updated this instrument based on the initial pilot data and conducted a second pilot study (Phase 5). In the following sections, we describe each of these pilot studies and the corresponding results.

Pilot Studies

Throughout the previous three phases, we primarily focused on gathering evidence of content validity for our instrument. In the following two phases, we conducted two separate pilot studies to obtain evidence of our instrument’s construct validity and reliability for measuring newer statistics instructors’ motivations to use group work. We created all figures and conducted our analysis using the statistical software R (R Core Team, 2020) and the *ggplot2* (*v.3.3.5*; Wickham, 2016), *semTools* (*v.0.5.*6; Jorgensen et al., 2022), and *lavaan* (*v.0.6.*11; Rossee, 2012) packages (Appendix I).

Phase 4: Pilot Study 1

For the first pilot study, we contacted faculty at four different universities (1 Western, 2 Midwestern, and 1 Eastern) to help facilitate the data collection process with newer statistics instructors. All the faculty we contacted worked with or supervised beginning instructors teaching statistics. Each was sent an email with a request to send an attached recruitment email to prospective participants at their institution. All four faculty contacts were willing to participate in the data collection process and sent the recruitment email to prospective participants. The recruitment email contained the Qualtrics link to our instrument and offered participants an opportunity to receive one of six $15 Amazon gift cards to be distributed to the 1st person and then every 10th person up until the 50th to complete the survey.

The instrument contained an initial consent form and screening questions to ensure that prospective participants were part of our intended population. Some screening questions included "Are you currently an instructor of a statistics course, or leading a statistics recitation/discussion section?" and "Are you currently a graduate student?" If a prospective participant was not teaching statistics, not using group work, or had more than three years of post-graduate teaching experience, they were excluded from our definition of newer statistics instructor and were immediately prompted to the end of the survey; they did not have an opportunity to respond to any of the motivational items. Prospective participants who were not a part of our intended population did not qualify for the monetary incentive.

Among the 20 prospective participants who answered the screening questions (Table 4.5), eight selected that they were either not a newer instructor or they did not use group work to teach statistics. One participant, who described themselves as part of the intended population, did not finish the instrument, leaving the last half blank. This suggests that their responses were not missing completely at random, so they were removed from the study.

Total number of prospective participants	**Prospective participants outside the intended population**	**Prospective participants within the intended population who gave incomplete responses**	**Prospective participants within the intended population who gave complete responses**
n = 20	n = 8	n = 1	n = 11

Table 4.5: Sample size breakdown for Phase 4.

To explore newer statistics instructors' responses on the instrument, we first assigned each response a numeric value: Strongly agree – 6; Agree – 5; Slightly agree – 4; Slightly disagree – 3; Disagree – 2; Strongly disagree – 1. None of the items were written to need reverse-coding, as these types of items are often misinterpreted by participants and reduce reliability of an instrument (Suárez-Alvarez et al., 2018). Next, we calculated summary statistics (Appendix J) across all items within a motivational construct for our 11 participants, revealing that, on average, using group work to teach statistics appeared to be a more self-determined behavior (intrinsic motivation: mean = 4.73; integrated regulation: mean = 4.93) than a less self-determined behavior (external regulation: mean = 3.93; amotivation: mean = 2.25) for them. We noticed that responses to intrinsic motivation items had the least amount of variability (*SD* = 0.89), while answers to external regulation items had the most amount of variability across our six response options (*SD* = 1.56).

For each motivational construct, we created line graphs of each participant's item responses to investigate evidence of construct validity (Appendix J). Several items displayed irregular patterns in participants' responses and were flagged for further investigation into their interpretation. For example, we flagged one integrated regulation item (IR2) because both Participant 1 and Participant 9 answered *Disagree*, while most of their responses to the other integrated regulation items were *Slightly agree* or higher (Figure 4.3). Upon further investigation, the irregular response patterns and data from the cognitive interviews suggested that for this item (IR2: *Using group work is essential when teaching statistics*), the word *essential* was being interpreted as *needed*. Because this item's interpretation did not align with our working definition of integrated regulation, we removed it from our instrument prior to Phase 5.

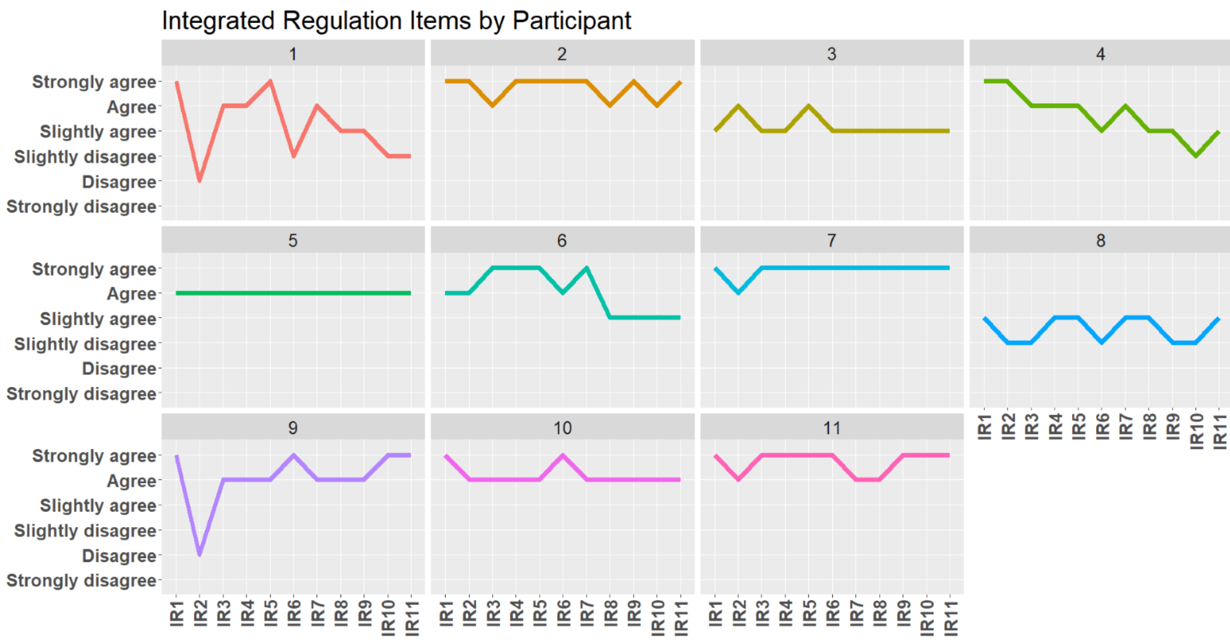


Figure 4.3: Line graphs of participant responses across integrated regulation items.

We flagged two amotivation items (AM2 and AM6). Similar to IR2, AM6 (*Group work is not necessary to teach statistics*) was interpreted more literally, with the participants interpreting the words *not necessary* as *not needed*. Because this item was not measuring the absence of intrinsic and extrinsic motivation as intended, it was removed from the instrument prior to Phase 5. In contrast, AM2 (*Typically, using group work to teach statistics ends up using more class time than it is worth*), which was also flagged, was considered to align with the construct amotivation and was not changed.

No intrinsic motivation items showed evidence of irregular responses, but the responses for our external regulation items (Figure 4.4) displayed consistent irregular patterns among most participants. For example, Participant 9 answered *Strongly disagree* for ER1, *Slightly agree* for

ER2, *Strongly agree* for ER3, and *Disagree* for ER4. Participants 6 and 8 had an equally extreme pattern of responses, and this evidence of irregular response patterns was corroborated by a lack of reliability evidence. Using Cronbach's alpha (Cronbach, 1951) for each set of items within the four latent motivational constructs to investigate reliability, there was evidence to suggest that the external regulation items were measuring more than one construct (7 items, $\hat{\alpha}$ = 0.569). In contrast, evidence from remaining Cronbach's alpha estimates suggested that intrinsic motivation items (8 items, $\hat{\alpha}$ = 0.952), integrated regulation items (11 items, $\hat{\alpha}$ = 0.929), and amotivation items (8 items, $\hat{\alpha}$ = 0.929) were all measuring their respectively intended construct. These estimates were calculated prior to removing the problematic items IR2 and AM6.

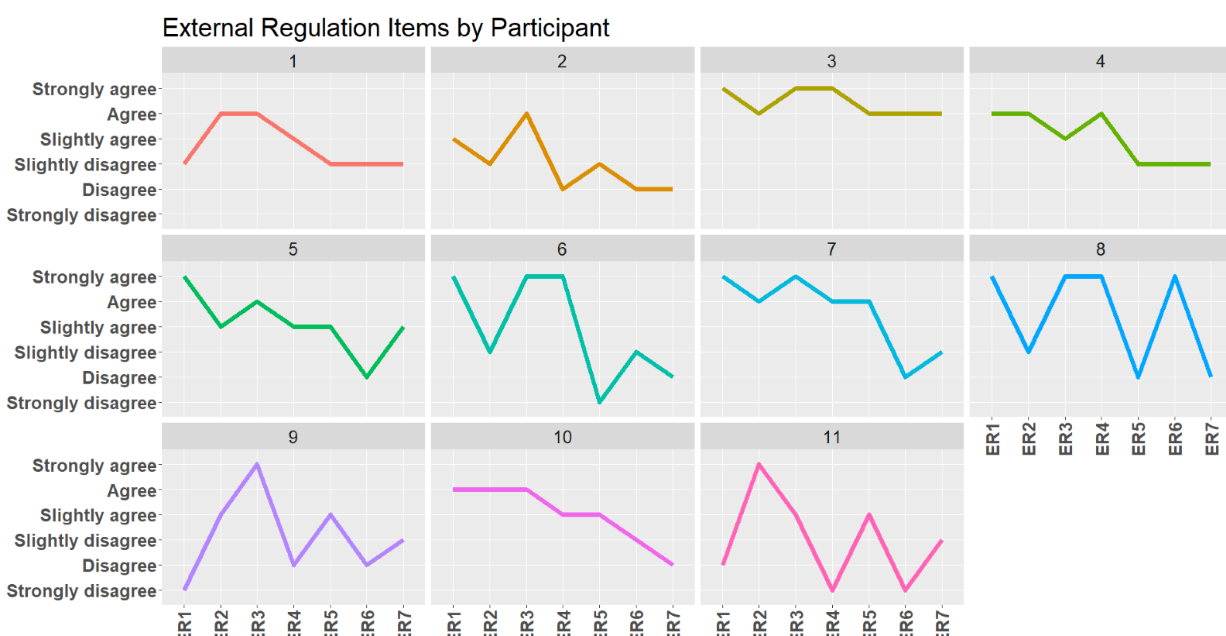


Figure 4.4: Line graphs of participant responses across external regulation items.

Based on the irregular response patterns and low evidence of reliability for the external regulation items, we further investigated each item individually. ER2, ER5, and ER7 were answered more similarly than the rest of the external regulation items and focused on using group work to earn an external reward (e.g., *One of the reasons I use group work is so that I will receive better reviews as an instructor*). The other four external regulation items focused on a different aspect of our working definition—pressure (e.g., *One of the reasons I use group work is because my colleagues or supervisors want me to use it to teach statistics*). Based on this evidence, we partitioned the external regulation items into two separate constructs: 1) external regulation – rewards and 2) external regulation – pressure. Cronbach's alpha estimates suggested items partitioned to rewards (3 items, $\hat{\alpha}$ = 0.809) and items partitioned to pressure (4 items, $\hat{\alpha}$ = 0.826) measure their respective intended construct. All these reliability estimates and bootstrapped estimated 95% confidence intervals to quantify the uncertainty around them can be found in Appendix J.

After deliberation, we concluded that more detailed information about a newer instructor's external regulation to use group work when teaching statistics may be valuable and decided to include both constructs on our instrument moving forward. We also wrote one more item for each of the newly defined constructs to collect additional information for measuring the *rewards* (RW) and *pressure* (PR) constructs.

Next, we calculated average inter-item correlations between sets of items across the five latent motivational constructs (amotivation (AM), external regulation – rewards (RW), external regulation – pressure (PR), integrated regulation (IR), and intrinsic motivation (IM)). For these and Phase 5's correlation calculations, we used Pearson's correlation. Although the nature of our data are ordinal, multiple studies have suggested that scales with five or more options are an ordinal approximation of a continuous variable and can be treated as such during analysis (Norman, 2010; Sullivan & Artino, 2013). Further, we chose to be consistent with the methodology common to other instruments when gathering evidence of validity and reliability (Holgado et al., 2010).

Inter-item correlations are created by calculating the correlations for every pair of items across different constructs and taking the mean (Piedmont, 2014). Motivations that classify more self-determined behavior (intrinsic motivation, integrated regulation) were positively related to one another (r = 0.61), while being negatively related to pressure (intrinsic motivation: r = -0.26, integrated regulation: r = -0.26) and showing little to no relationship with rewards (intrinsic motivation: r = -0.05, integrated regulation: r = -0.10). Amotivation showed a strong negative relationship with intrinsic motivation (r = -0.55) and integrated regulation (r = -0.52) and little to no relationship with pressure (r = 0.19) and rewards (r = -0.07). Based on these results, there is initial evidence of construct validity—that the responses to our items are consistent with the relationships outlined in SDT (Appendix J).

After making all changes based on evidence collected in Phase 4, our instrument contained eight intrinsic motivation items, eight integrated regulation items, four rewards items, five pressure items, and seven amotivation items. These collective items were used in Phase 5.

Phase 5: Pilot Study 2

Data Collection The purpose of Phase 5 was to continue assessing evidence of construct validity and reliability for our instrument, but with a larger sample size. To obtain a larger sample size, we used multiple approaches for recruitment. First, we contacted and posted a recruitment statement across three statistics and statistics education listservs: American Statistical Association (ASA) Section on Statistics and Data Science Education; Consortium for the Advancement of Undergraduate Statistics Education (CAUSE); and Isolated Statisticians. Second, we composed a list of 58 Statistics Ph.D.-granting universities and contacted individuals from each university who worked with newer statistics instructors or were graduate program directors. Each was sent an email with a request to send an attached recruitment email to prospective participants at their institution. Individuals at 15 of the 58 universities responded and agreed to participate in the data collection process. Additionally, 12 experts who participated in previous phases of data collection were re-contacted to send an attached recruitment email to prospective participants, and 10 agreed to participate. Newer statistics instructors were also directly contacted and invited to participate in the study.

We used the same consent form and screening questions as in Phase 4 to ensure we were collecting data on newer statistics instructors. Among the 215 prospective participants who filled out the screening questions, 157 selected that they were either not a newer instructor or they did not use group work to teach statistics (Table 4.6), so they were immediately prompted to the end of the survey and did not have an opportunity to respond to any of the motivational items. Six participants who described themselves as being a part of the intended population did not finish the survey. For all six, there was sufficient evidence that their missing responses followed a distinct pattern and were not missing completely at random, so they were removed before analysis. This left us with 58 participants for the analysis.

Number of Prospective Participants	**Prospective participants outside the intended population**	**Prospective participants within the intended population who gave incomplete responses**	**Prospective participants outside the intended population who gave complete responses**
n = 215	*n* = 157	*n* = 6	*n* = 58

Table 4.6: Sample size breakdown for Phase 5.

Data Analysis We used multiple forms of data analysis to gather evidence of construct validity and reliability for our instrument. We first created a correlation matrix and calculated average inter-item correlations to initially explore evidence of construct validity. Since we were able to achieve a larger sample size in Phase 5 than in Phase 4, we calculated coefficient omega instead of Cronbach's alpha. Cronbach's alpha assumes tau-equivalence, which may not be practical and can underestimate reliability coefficients; coefficient omega does not make that same assumption, but requires a larger sample size (Teo & Fan, 2013). We then continued to assess construct validity by fitting a confirmatory factor analysis (CFA) model; CFA estimates how "good" our items are by estimating the relationship between our items and their associated motivational construct. For our model to be identifiable, some parameters must be fixed. Common methodology when fitting a CFA model is to fix one factor loading within each construct to be equal to one, so the model is identifiable and the factor loading interpretations are clearer. This approach is justifiable when one item is known to be a "goldstandard" indicator of a latent factor of interest (Little, 2013). However, this type of assumption is inappropriate to make during this exploratory phase of instrument development. At this point, we have no evidence to suggest that any one item we've created is perfectly related to its latent motivational construct. Therefore, when identifying our model, we instead fixed each factor's latent variance equal to one, allowing for all factor loadings to be estimated and explored; this model specification also uses an oblique rotation which assumes that the motivational constructs are correlated, adhering to SDT's claim that multiple motivations can be experienced at the same time and are not mutually exclusive.

Results Summary statistics revealed that the use of group work for teaching statistics was characterized, on average, for participants as more of a self-determined behavior (intrinsic motivation: mean = 4.72; integrated regulation: mean = 5.06) than a lesser self-determined behavior (pressure: mean = 3.45; rewards: mean = 3.16; amotivation: mean = 2.16). Average standard deviations of item responses ranged between 0.81 and 1.55 across our five motivational constructs. Three of the five constructs had a standard deviation less than one (intrinsic motivation, integrated regulation, amotivation). The average standard deviations across the different sets of items revealed that integrated regulation items tended to be answered the most similarly (*SD*= 0.81). IR1 and IR5 yielded extremely low standard deviations (*SD*= 0.67 and *SD*= 0.68, respectively); neither item had *Strongly disagree, Disagree, or Slightly disagree* selected as a response. Histograms of response patterns for each question are found in Appendix J.

We constructed a Pearson correlation matrix (Figure 4.5) to assess the relationships among items both within and across constructs. Overall, we observed patterns across constructs consistent with what we would expect under the framework of SDT. There is a clear negative relationship between items that represent amotivation and more self-determined types of motivation (intrinsic motivation, integrated regulation). In addition, there appears to be little to no relationship between less self-determined external regulations (pressure, rewards) and amotivation. Intrinsic motivation and integrated regulation, constructs that measure more self-determined behavior, appear to be extremely related. Results primarily reinforced our initial evidence of validity towards measuring the five intended constructs.

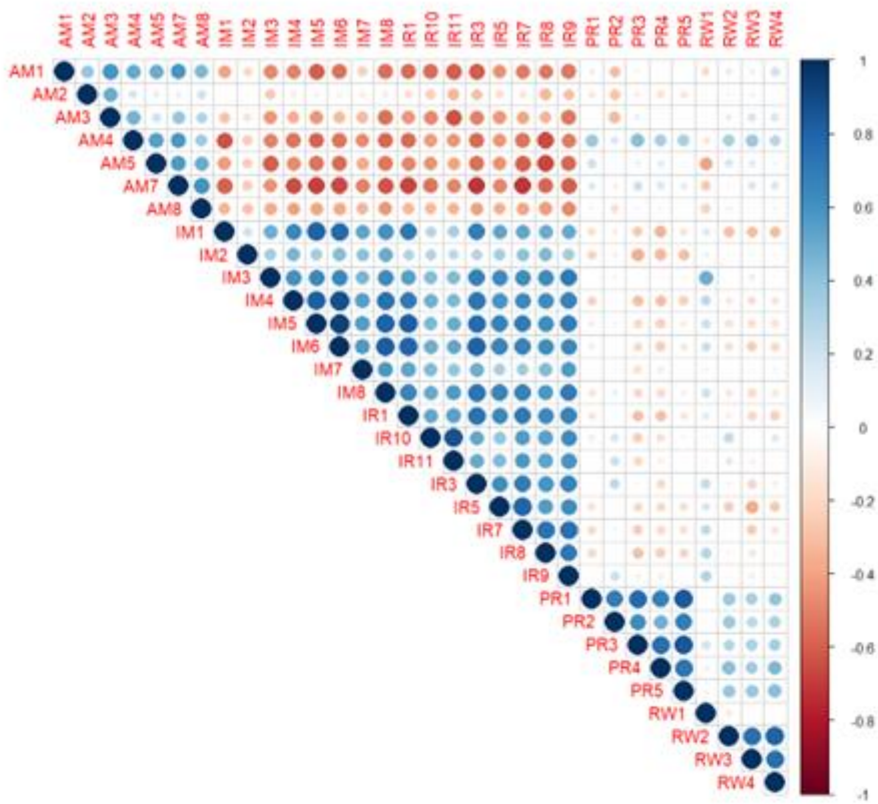


Figure 4.5: Pearson correlation matrix for 32 items in Phase 5. The size and color of the dot represent the strength and direction of the relationship, respectively. Larger blue dots represent stronger positive correlations while larger red dots represent stronger negative correlations between items.

The estimated relationships among most items within each construct were fairly strong, but two items appeared to be unrelated to other items that intended to measure the same construct. AM2 (*Typically, using group work to teach statistics ends up using more class time than it is worth*) had a weak positive relationship with four of the six other amotivation items, while RW1 (*I use group work because it typically makes teaching statistics easier*) showed evidence of no relationship with all other items intended to measure extrinsic motivation stemming from rewards.

Coefficient omega calculations (McDonald, 1999) suggest that intrinsic motivation items (8 items, $\Omega = 0.925$), integrated regulation items (8 items, $\Omega = 0.925$), pressure items (5 items, $\Omega = 0.931$), reward items (4 items, $\Omega = 0.824$) and amotivation items (7 items, $\Omega = 0.811$) all measure their respectively intended motivational construct. Maximum likelihood estimated 95% confidence intervals to better quantify the uncertainty around these estimates can be found in Appendix J.

To continue gathering evidence of construct validity, we proposed a five-factor CFA model to assess the relationship between intrinsic motivation, integrated regulation, pressure, rewards, amotivation, and their respective items. Additionally, this model estimated the relationships between our five latent motivational constructs (Figure 4.6).

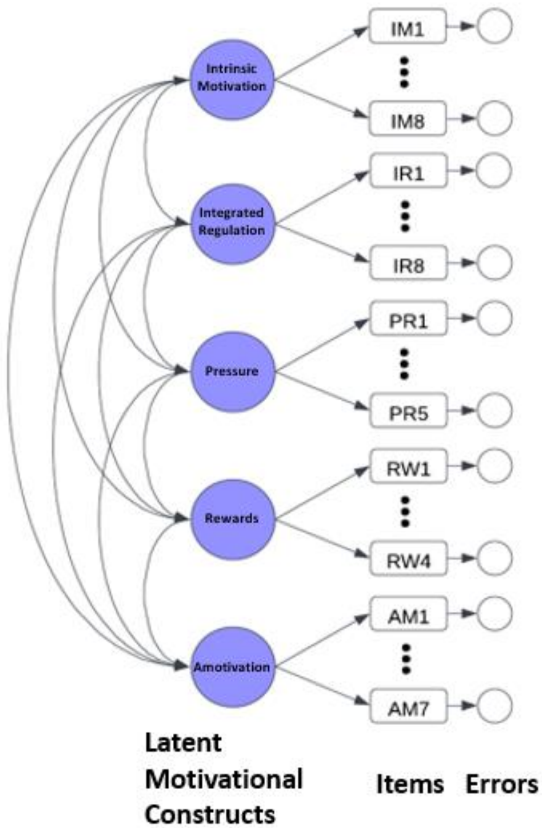


Figure 4.6: Latent structure for measurement model fit to measure motivation. Solid blue circles represent the five motivational constructs. Arrows connecting the blue circles represent the estimated correlation shared between each pair of motivational constructs. The straight protruding arrows connect the constructs to their corresponding items. The boxes represent the items. Protruding arrows from the item boxes connect to each item's error variance.

According to Hair et al. (2010), acceptable factor loading estimates should be higher than 0.5. Results from CFA found that all but three items had standardized factor loadings higher than this value (Table 4.7). These results suggest that the proportion of variance explained by each of these item's respective latent factor is small and that there is little relationship between the motivational latent factor and each corresponding item.

Item	Item Text	Estimated Standardized Factor Loading	Standard Error	P-value
IM1	I find it interesting to use group work when teaching statistics	0.792	0.094	< 0.001
IM2	***I have a desire to continue learning about how to use group work when teaching statistics***	***0.435***	***0.105***	***0.001***
IM3	My favorite way to teach statistics is through the use of group work	0.688	0.142	< 0.001
IM4	I get excited when using group work to teach statistics	0.880	0.107	< 0.001
IM5	I take pleasure in using group work when teaching statistics	0.948	0.098	< 0.001
IM6	I find it enjoyable to use group work when teaching statistics	0.966	0.093	< 0.001
IM7	Learning how to use group work has been a process I enjoy	0.604	0.123	< 0.001
IM8	I have fun when using group work to teach statistics	0.854	0.093	< 0.001
IR1	Using group work helps me achieve my goals as an instructor when teaching statistics	0.838	0.071	< 0.001
IR3	Typically, using group work to teach statistics is rewarding for me as an instructor	0.841	0.087	< 0.001
IR5	Typically, I find using group work valuable for my students when teaching statistics	0.740	0.077	< 0.001
IR7	Typically, I find using group work worthwhile for my students when teaching statistics	0.855	0.077	< 0.001
IR8	I am committed to consistently using group work when teaching statistics	0.788	0.108	< 0.001
IR9	It is important to me to use group work when teaching statistics	0.835	0.090	< 0.001
IR10	Using group work to teach statistics reflects what I believe in as an instructor	0.681	0.100	< 0.001
IR11	Using group work is consistent with my values as a statistics instructor	0.684	0.102	< 0.001
RW1	***I use group work because it typically makes teaching statistics easier***	***-0.065***	***0.160***	***0.636***

Table 4.7: Standardized factor loadings, standard errors, and p-values from CFA results. Three items identified with low standardized factor loadings are bolded and italicized.

Table 4.7 Continued

RW2	One of the reasons I use group work when teaching statistics is to be viewed as a better instructor	0.899	0.141	< 0.001
RW3	One of the reasons I use group work is so that I will receive better reviews as an instructor	0.828	0.133	< 0.001
RW4	One of the reasons I use group work when teaching statistics is to be evaluated as a better instructor by others	0.901	0.157	< 0.001
PR1	One of the reasons I use group work is because others want me to use it to teach statistics	0.883	0.165	< 0.001
PR2	One of the reasons I use group work to teach statistics is because it is recommended to me by others	0.724	0.162	< 0.001
PR3	One of the reasons I use group work to teach statistics is because I am told to	0.896	0.179	< 0.001
PR4	One of the reasons I use group work to teach statistics is because I am pressured by others to do so	0.783	0.159	< 0.001
PR5	One of the reasons I use group work to teach statistics is because others expect me to do so	0.939	0.153	< 0.001
AM1	I do not see what value using group work brings me when teaching statistics	0.745	0.085	< 0.001
AM2	***Typically, using group work to teach statistics ends up using more class time than it is worth***	***0.304***	***0.164***	***0.023***
AM3	Using group work to teach statistics is overrated as a teaching technique	0.587	0.132	< 0.001
AM4	I do not care much about using group work when teaching statistics	0.723	0.107	< 0.001
AM5	I typically choose not to use group work when teaching statistics	0.692	0.102	< 0.001
AM7	I do not like using group work to teach statistics	0.810	0.109	< 0.001
AM8	I do not find it useful to use group work when teaching statistics	0.577	0.119	< 0.001

Table 4.8: Standardized factor loadings, standard errors, and p-values from CFA results. Three items identified with low standardized factor loadings are bolded and italicized.

The estimated correlations between the latent traits show that motivational constructs that categorize the use of group work as more self-determined are more related to each other than motivational constructs that categorize lesser self-determined behavior. As expected under the lens of SDT, we observe a strong positive relationship between intrinsic motivation and integrated regulation ($r = 0.891$). Both intrinsic motivation and integrated regulation showed little to no evidence of a relationship with pressure (intrinsic motivation: $r = -0.181$, integrated regulation: $r = -0.166$). Intrinsic showed some evidence of a negative relationship with rewards ($r = -0.213$) while there was little evidence of a relationship between integrated regulation and rewards ($r = -0.124$). Amotivation showed a strong negative relationship with intrinsic motivation ($r = -0.809$) and integrated regulation ($r = -0.903$), while showing some evidence of a positive relationship with rewards ($r = 0.272$) and little to no evidence with pressure ($r = 0.171$). The chi-square fit statistic and root mean square error of approximation (RMSEA) that are often used to assess model fit can be found in Appendix J.

Discussion and Conclusions

This research investigated how to measure motivations for newer statistics instructors' use of active learning through the framework of SDT. In Phases 1 through 3, we gathered evidence of content validity through the solicitation of expert opinion and cognitive interviews. In our first pilot study, we observed initial evidence of construct validity and reliability among three of our four latent motivational constructs. The lack of validity and reliability evidence for external regulation led to partitioning external regulation into two separate constructs: pressure and rewards.

From the CFA, we discovered three items (RW1, AM2, and IMW) had a weak relationship with their intended construct (Table 4.7). This has implications for moving forward with instrument development. Comparing RW1 (*I use group work because it typically makes teaching statistics easier*) to other reward items, we observed that this item was worded in a way that the reward driving the use of active learning, making teaching easier, was more abstract and subjective than the others. Upon deliberation, we determined that newer statistics instructors using group work may conclude that this statement is false, and respondents could conclude that group work makes it harder to teach statistics due to factors such as increased lesson planning and class time. Thus, it becomes unclear for participants if they should agree or disagree with this item. In future phases of survey development, we plan to replace this item.

The second problematic item, AM2 (*Typically, using group work to teach statistics ends up using more class time than it's worth*) may focus on characteristics of instruction rather than measuring motivation to use group work. This item is in reference to *time*, having newer statistics instructors compare how much time group work takes to their values about the technique's worth. It is plausible that an instructor who is amotivated to use active learning may disagree with this item because when they use group work, it does not take up much class time. We plan to further investigate how this item is being interpreted prior to future pilot studies.

We hypothesize the low standardized factor loading for the third item, IM2 (*I have a desire to continue learning about how to use group work when teaching statistics*), is the result of having two sub-populations, GSIs (n = 28) and non-GSIs (n = 30), within our population of newer statistics instructors. Across these two sub-populations, non-GSIs responded *Strongly agree* to this item more often than GSIs. This difference could be attributed to the terminal nature of teaching for some GSIs. For example, a GSI may be intrinsically motivated to use group work when teaching statistics, but they might not desire to learn more about using group work if they are seeking non-academic positions. Thus, there is initial evidence that this item may not appropriately measure intrinsic motivation for GSIs, which needs to be addressed prior to future administrations.

Given this discovery, we further investigated potential differences in response patterns between GSIs and non-GSIs by calculating the relative frequencies of each sub-population's item responses (Appendix J). This investigation revealed that, for our sample, GSIs were more likely to be extrinsically motivated by pressure to use active learning than newer instructors who were not GSIs. This preliminary finding may hint at why the adoption to use active learning is even more difficult for GSIs, who are often in their very beginning stages of teaching (Justice, 2020).

Overall, this study gathered initial evidence of construct validity, content validity, and reliability. This evidence helps lay the foundation to continue gathering evidence of validity and reliability towards developing this instrument that intends to measure each of the five motivational constructs. Development of this instrument will inform future research on statistics instructors' motivations to use active learning and sets the groundwork to explore how to promote the use of group work as a more self-determined behavior when teaching statistics.

Limitations

One limitation of this study is sample size. Due to the small sample size, assumptions for our five-factor CFA model may have been violated, including multivariate normality, or that our data are normally distributed. The lack of variability among, and clear patterns within item responses (e.g., left-skewed item response distributions for more self-determined types of constructs; right-skewed item response distributions for amotivation) may suggest that the assumption of multivariate normality of our responses is unreasonable. With our relatively small sample size, it is possible that the population distribution of responses is approximately normal and we just observed potentially unusual responses due to our smaller sample size. It is also possible that despite the varied recruitment approaches we used, our sampling procedures may have resulted in a biased sample, leading to data on newer statistics instructors who have motivations that characterize their use of active learning as more self-determined. Consequently, these may not be representative of all newer statistics instructors who use group work. It is plausible that participants willing to fill out the instrument are more likely to show more self-determined types of motivations for using group work when teaching statistics than those who were not willing to fill out the instrument, resulting in the item response distributions we observed. Our small sample size also made it inappropriate to conduct any inferential investigations, including investigations of model fit. In the future, formal inferential investigations should be conducted to further support this instrument's development.

Another limitation of this research is that we do not measure introjection or identified regulation on our instrument. Additional studies may consider the inclusion of these constructs to further understand newer instructors' situational motivation to use group work when teaching statistics. It is acknowledged that, for a more complete understanding of changes in instructor motivations, situational motivation, contextual motivation, and global motivation should be studied simultaneously (Vallerand, 2007). Studying and distinguishing between these three levels allows for a more holistic understanding of why newer statistics instructors engage in group work than just studying situational motivation alone.

Future Research

Although our initial analysis suggested that our items measured five motivational constructs, there was evidence to suggest that three items may be problematic and their use should be reexamined prior to additional rounds of instrument validation; either these three items need to be better aligned with their respective motivational construct or removed from the instrument.

After addressing concerns about these items, future research pertaining to this instrument's development must include additional rounds of refinement with larger sample sizes to inferentially investigate model fit and measurement invariance across sub-populations. This includes testing for measurement invariance between our two sub-populations of instructors (GSIs and non-GSIs) to ensure that motivation is being measured similarly for each. Additionally, we plan to compare relationships between the constructs using both Pearson's correlation and polychoric correlations. Although the use of Pearson's correlation is justifiable, polychoric correlations respect the ordinal nature of these types of data and may produce a better measurement model to assess evidence of validity and reliability (Holgado et al., 2010)

Group work is not the only form of active learning on which experts prioritized measuring motivations, so future studies should investigate if and how the psychometric properties of this instrument change when adapting the context to a different type of active learning technique. Similarly, other disciplines such as Mathematics, recommend the use of active learning when teaching (CBMS, 2016). Future studies should investigate if and how the psychometric properties of this instrument change when adapting the context for a different discipline.

Future research should also investigate evidence for criterion validity. Evidence for criterion validity will ensure that the results from this instrument indeed align with predicting newer statistics instructors' use of group work as a self-determined behavior, better ensuring that results from this instrument support future studies' investigations of motivation.

The development of an instrument with validity and reliability evidence for measuring newer statistics instructors' motivations helps lay the groundwork to better understand why newer statistics instructors use group work. This information will uncover how self-determined the use of group work is when one teaches statistics and will inform how to better foster the use of group work as a more self-determined behavior for newer statistics instructors. Understanding how to foster the use of group work as a more self-determined behavior will better align current instructional practices of newer statistics instructors with current national recommendations to use group work when teaching statistics.

Acknowledgements

We thank each expert who filled out surveys to inform this work, the six GSIs who participated in the cognitive interviews, and all respondents who filled out our instrument. The time invested into helping with the development of this instrument is appreciated. We also thank Laura Hildreth, Nicola Justice, and Mark Greenwood for their insightful comments and conversations throughout the different phases of this study.

CONCLUSION

National recommendations call for the use of active learning in statistics classrooms. Yet, the adoption of active learning techniques remains challenging, especially for newer statistics instructors who may not be ready to use such advanced techniques (Beisiegel, 2019; Justice, 2020). This has created a need to better understand newer statistics instructors' knowledge about and emotions and motivations towards active learning. To address this need, we identified challenges graduate student instructors (GSIs) face when using active learning and investigated how GSIs break through these challenges.

Research suggests that instructors' knowledge about and emotions towards using active learning can promote or inhibit their use of such techniques (Finelli et al., 2014; Keller, Chang, Becker, Goetz, & Frenzel, 2014). While multiple studies have identified lack of knowledge and negative emotions as challenges for using active learning, little research has investigated when and how instructors overcome these challenges. My experiences as a GSI, working with GSIs, and wanting to help support GSIs' best practices influenced my research direction and inspired me to research this phenomenon.

We used purposeful sampling of two GSIs—one that identified as a "facilitator" and one that identified as a "lecturer"—and we collected survey data, video data, and semi-structured interview data to investigate their knowledge about, emotions towards, and use of active learning across multiple semesters. These data uncovered that both GSIs were in survival mode when first starting to teach, predominantly relying on lecture-based methods, before experiencing breakthroughs with their knowledge, emotions, and ultimately their use of active learning in subsequent semesters. Common themes across each GSI's breakthroughs included the need to feel confident in and be challenged by their course structure, and to align their knowledge about and emotions towards active learning before implementing it in their classroom. These results substantiate the call to support instructors' use of active learning with curricula that helps instructors facilitate such techniques (Haack, 2008) and highlights the importance emotions can play in trying to support the use of active learning in classrooms. Emotions can promote or deter instructors' practices (Burić & Frenzel, 2020). Thus, it isn't enough to educate and increase instructors' knowledge about active learning, and instead, more emphasis should be placed on their emotional responses when being introduced to such techniques.

This research adds to the existing literature in two main ways. First, this longitudinal study sheds light on how GSIs break through identified challenges over time. Investigating context around how and when these breakthroughs occur leads to insightful discussion about how to better facilitate these breakthroughs with active learning sooner. Highlighting the alignment of knowledge about and emotions towards active learning as a key to facilitate breakthroughs calls on support systems to consider an additional layer of complexity when designing and deciding best ways to support the adoption of active learning. In addition, this research emphasizes that longitudinal support is necessary and could help foster breakthroughs with active learning that happen in later semesters.

Secondly, this research sets the foundation to investigate the use of active learning as a self-determined behavior. To do this, we reflected on how we could measure "why" newer statistics instructors were using active learning. Motivation, the "why" of behavior, plays a critical role in behavioral change and is a key component to align statistics instructors' practices with national recommendations to use active learning. However, there is currently little research around motivation in the context of newer statistics instructors and active learning. To pursue this investigation, we needed to first establish a comprehensive understanding of motivation and justify a framework for studying newer statistics instructors' motivations to use active learning.

Through a review of current and common motivational frameworks, we chose to study motivation through the lens of self-determination theory (SDT). We argue that this motivational framework provides unique and practical information that can inform how to support statistics instructors' use of active learning as a self-determined behavior. If using active learning when teaching statistics is self-determined, statistics instructors will be more likely to persist and engage with its use. The three different types of motivation and six subsequent regulatory styles theorized by SDT describe why an instructor is using active learning and classify the use of active learning on a continuum of self-determined behavior.

We also related SDT to self-regulation learning (SRL) theory, presenting a model for fostering active learning as a more self-determined behavior. SDT emphasizes that a behavior becomes more self-determined if one's psychosocial needs for autonomy, competence, and relatedness are fulfilled. SRL theory presents a cycle that helps fulfill these psychosocial needs when one is learning about how to use active learning techniques. Through their common connection to psychosocial needs, these two theories outline a model for fostering self-determined behavior that could inform support systems on how to better encourage effective and continued use of active learning in statistics classrooms.

Although much is understood about motivation, less is known about how to foster active learning as more a more self-determined behavior. Because there is no known method or instrument to quantify newer statistics instructors' motivations to use active learning, we developed a multi-phase study and used SDT to create an instrument that measures intrinsic motivation, integrated regulation, external regulation, and amotivation and established initial evidence of its validity and reliability. The first three phases of the study included the solicitation of expert opinion and cognitive interviews with newer statistics instructors. Data from experts and the cognitive interviews uncovered potential interpretability issues and allowed us to assess each item's relationship to its intended motivational construct. These phases established evidence of content validity among our 34 items created. Two pilot studies were then conducted with newer statistics instructors who use group work when teaching statistics. Evidence from the first pilot study suggested that our crafted items that measured external regulation were not reliable. Additional investigation suggested that two separate constructs within the definition of external regulation were being measured: pressure and rewards. Because this level of detail may be practically important for those using the instrument, we decided to separate external regulation into two separate constructs and conducted an additional pilot with a larger sample size.

Confirmatory factor analysis and reliability estimates provided initial evidence of reliability and validity, suggesting that this instrument may help inform future studies' investigations of motivation. There was evidence of three problematic items that need to be updated prior to future pilots. Future studies will be conducted to continue obtaining evidence of validity and reliability towards measuring motivation for newer statistics instructors' use of group work.

Existing research has identified active learning's place in statistics classrooms and the challenges around its implementation. Our research expands this understanding and provides a newly developed instrument to help inform future studies on how to facilitate the adoption of active learning in statistics classrooms.

Directions for Future Research

Student benefits—such as improved retention, increased critical thinking, and the development of collaborative skills—are just some of many reasons why active learning has cemented its place in statistics education (Freeman et al., 2014). However, national recommendations to incorporate active learning techniques into statistics classrooms, and the challenges in doing so, leave a need to better understand the relationship between newer statistics instructors and active learning.

The need for professional development programs to help facilitate the adoption of active learning are a growing necessity (Pfund et al., 2009). While models for professional development programs to promote active learning are in place, recommendations on how to support newer statistics instructors use of active learning are often not connected to motivation. Research must outline the relationships between professional development decisions to support the use of active learning and how they impact an instructor's motivations when using active learning. This will create a more comprehensive understanding how developmental decisions support the use of active learning as a self-determined behavior. An example of this research could be presented from the combination of our previous studies in Chapter 2 and Chapter 4. When Max first started using active learning, was it considered a self-determined behavior? What types of support did Max experience that fostered the use of active learning as more self-determined? If these types of support would have been given earlier, would we have seen breakthroughs with active learning in earlier semesters?

With active learning here to stay, research needs to continue to push for a more complete understanding on how to foster early adoption of such techniques in statistics classrooms. We hope this research promotes conversation on how newer statistics instructors can break through with active learning and use it during their early development with proper support. Additionally, we want to emphasize motivation’s role in facilitating change and stress the importance to understand the use of active learning through this lens. We hope this instrument, and this emphasis, can inform future studies on how to better support the use of active learning for newer statistics instructors.

REFERENCES CITED

Allsop, J., Young, S. J., Nelson, E. J., Piatt, J., & Knapp, D. (2020). Examining the benefits associated with implementing an active learning classroom among undergraduate students. *International Journal of Teaching and Learning in Higher Education, 32*(3), 418-426.

American Association of University Professors. (2018). Data Snapshot: Contingent Faculty in US Higher Ed. Retrieved from https://www.aaup.org/sites/default/files/10112018 Data Snapshot Tenure.pdf

American Statistical Association. (n.d.). *Statistics is specifically required in some majors, while it is a quantitative methods requirement in others*. Retrieved from College Majors Requiring Statistics. Retrieved from https://www.amstat.org/asa/files/pdfs/EDU-CollegeMajorsFlyer.pdf

Ames, H., Glenton, C., & Lewin, S. (2019). Purposive sampling in a qualitative evidence synthesis: A worked example from a synthesis on parental perceptions of vaccination communication. *BMC Medical Research Methodology, 19*(26), 1-9.

AMSPP. (2021). *Islands in schools project*. Retrieved from Australian Government Department of Education. Retrieved from https://sites.google.com/site/islandsinschoolsprojectwebsite/home

Andrews, T. C., Auerbach, A. J., & Grant, E. F. (2019). Exploring the relationship between teacher knowledge and active-learning implementation in large college biology courses. *CBE—Life Sciences Education, 18*, ar48.

Apkarian, N., Henderson, C., Stains, M., Raker, J., Johnson, & Dancy, M. (2021). What really impacts the use of active learning in undergraduate STEM education? Results from a national survey of chemistry, mathematics, and physics instructors. *PLOS ONE, 16*(2), 1-15.

Auerbach, A. J., Higgins, M., Brickman, P., & Andrews, T. C. (2018). Teacher knowledge for active-learning instruction: Expert–novice comparison reveals differences. *CBE—Life Sciences Education, 17*, ar12.

Ausubel, D. P. (1968). *The psychology of meaningful verbal learning*. Grune and Stratton.

Beatty, P., & Willis, G. (2009). Research synthesis: The practice of cognitive interviewing. *Public Opinion Quarterly*, *71*(2), 1-25.

Beisiegel, M. (2019). Teacher interrupted: How mathematics graduate teaching assistants (don't) learn about teaching. *Proceedings of the 43rd Conference of the International Group for the Psychology of Mathematics Education, 2.*

Blair, R. M., Kirkman, E. E., Maxwell, J. W., & Society, A. M. (2013). *Statistical abstract of undergraduate programs in the mathematical sciences in the United States: Fall 2010 CBMS survey.* American Mathematical Society.

Blais, M., Vallerand, R., Brière, N., Gagnon, A., & Pelletier, L. (1990). Significance, structure, and gender differences in life domains of college students. *Sex Roles*, *22*, 199-212.

Boler, M., & Greene, M. (1999). *Feeling power: Emotions and education.* Routledge.

Bonwell, C. C., & Eison, J. A. (1991). *Active learning: creating excitement in the classroom.* George Washington University, ERIC Clearinghouse on Higher Education.

Brenner, C. (2022). Self-regulated learning, self-determination theory and teacher candidates' development of competency-based teaching practices. *Smart Learning Environments, 9*.

Burić, I., & Frenzel, C. A. (2020). Teacher emotional labour, instructional strategies, and students' academic engagement: A multilevel analysis. *Teachers and Teaching: Theory and Practice*, *21*(5), 335-352.

Carver, R. H., Everson, M., Gabrosek, J., Horton, N. J., Lock, R. H., Mocko, M., . . . Wood, B. (2016). Guidelines for Assessment and Instruction in Statistics Education (GAISE) College Report 2016.

Center for Research on Learning and Teaching. (n.d.). *How can you incorporate active learning into your classroom?* Retrieved from https://crlt.umich.edu/sites/default/files/Active_Learning_Continuum_CRLT.pdf

Chang, L. (1994). A psychometric evaluation of 4-point and 6-point likert-type scales in relation to reliability and validity. *Applied Psychological Measurement*, *18*, 205-215.

Chyung, S. Y., Roberts, K., Swanson, L., & Hankinson, A. (2017). Evidence-based survey design: The use of a midpoint on the likert scale. *Performance Improvement, 56*, 15-23.

Conference Board of Mathematical Sciences. (2016). Active learning in post-secondary mathematics education. In *Conference Board of the Mathematical Sciences*.

Cotner, S., Loper, J., Walker, J., & Brooks, C. D. (2013). Research and teaching: "It's not you, it's the room"--Are the high-tech, active learning classrooms worth it? *Journal of College Science Teaching, 42*(6), 82-88.

Creswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches.* Thousand Oaks, CA: Sage.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16,* 297-334

Davidian, M., & Louis Thomas, A. (2012). Why statistics? *Science, 335*(6077), 12.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic Motivation and Self-Determination in Human Behavior*. Berlin: Springer Science & Business Media.

Deshler, J. M., Hauk, S., & Speer, N. (2015). Professional development in teaching for mathematics graduate students. *Notices of the American Mathematical Society, 62*(6), 638-643.

Deslauriers, L., McCarty, L., Miller, K., Callaghan, K., & Kestin, G. (2019). Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom. *Proceedings of the National Academy of Sciences, 116*, 201821936. doi:10.1073/pnas.1821936116

El-Ghoroury, N. H., Galper, D., Sawaqdeh, A., & Bufka, F. L. (2012). Stress, coping, and barriers to wellness among psychology graduate students. *Training and Education in Professional Psychology, 6*(2), 122-134.

Ellis, J. (2014). Preparing future professors: Highlighting the importance of graduate student professional development Programs in calculus instruction. *Proceedings of the Joint Meeting of PME 38 and PME-NA, 36*, 9-16.

Evelein, F., Korthagen, F., & Brekelmans, M. (2008). Fulfillment of the basic psychological needs of student teachers during their first teaching experiences. *Teaching and Teacher Education, 24*(5), 1137-1148.

Fernet, C., Senécal, C., Guay, F., Marsh, H., & Dowson, M. (2008). The work tasks motivation scale for teachers. *Journal of Career Assessment, 16*, 256-279.

Finelli, C., Daly, S., & Richardson, K. (2014). Bridging the research-to-practice gap: Designing an institutional change plan using local evidence. *Journal of Engineering Education, 103*(2), 331-361.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences, 111*(23), 8410–8415.

Friedman, J. (2017). *Universities Where TAs Teach the Most Classes*. World Report. Retrieved from https://www.usnews.com/education/best-colleges/the-short-list-college/articles/2017-02-21/10-universities-where-tas-teach-the-most-classes

Funke, J. (2017). How Much Knowledge Is Necessary for Action? In Klaus Tschira Symposia, *Knowledge and Action*. Springer. doi:10.1007/978-3-319-44588-5_6

Garfield, J., & Everson, M. (2009). Preparing teachers of statistics: A graduate course for future teachers. *Journal of Statistics Education*, *17*(2), 223-227.

Garland, R. (1991). The mid-point on a rating scale: Is it desirable? *Marketing Bulletin*, *2*, 66-70.

Gilar, R., Pozo-Rico, T., Sánchez, B., & Castejon, J. L. (2019). Can emotional intelligence be improved? A randomized experimental study of a business-oriented eI training program for senior managers. *PLOS ONE*, *14*, e0224254.

Grasha, A. F. (1994). A matter of style: The teacher as expert, formal authority, personal model, facilitator, and delegator. *College Teaching*, *42*(4), 142–149.

Gross, D., Pietri, S. E., Anderson, G., Moyano-Camihort, K., & Graham, J. M. (2015). Increased preclass preparation underlies student outcome improvement in the flipped classroom. *CBE—Life Sciences Education*, *14*(4), 1-8.

Guay, F., Vallerand, R., & Blanchard, C. (2000). On the assessment of situational intrinsic and extrinsic motivation: The situational motivation scale (SIMS). *Motivation and Emotion*, *24*(3), 175-213.

Haack, K. (2008). UN studies and the curriculum as an active learning tool. *International Studies Perspectives*, *9*(4), 395-410.

Hagger, M., Chatzisarantis, N., Hein, V., Pihu, M., Soós I., & Istvan, K. (2007). The perceived autonomy support scale for exercise settings (PASSES): Development, validity, and cross-cultural invariance in young people. *Psychology of Sport and Exercise*, *8*, 632-653.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Upper Saddle River, NJ: Person Prentice Hall.

Haladyna, T. M., & Rodriguez, M. C. (2013). *Developing and validating test items* (1st ed.). New York: Routledge.

Han, J., & Yin, H. (2016). Teacher motivation: Definition, research development and implications for teachers. *Cogent Education, 3,* 1217819.

Hancock, P., & Volante, W. (2020). Quantifying the qualities of language. *PLOS ONE, 15,* e0232198.

Hargreaves, A. (2005). Educational change takes ages: Life, career and generational factors in teachers' emotional responses to educational change. *Teaching and Teacher Education*, *21*, 967-983.

Harmin, M., & Melanie, T. (2006). *Inspiring active learning: A complete handbook for today's teachers.* Association for Supervision and Curriculum Development.

Heim, S., & Keil, A. (2017). Too much information, too little time: How the brain separates important from unimportant things in our fast-paced media world. *Frontiers for Young Minds, 5*(23).

Henderson, C., Dancy, M., & Niewiadomska-Bugaj, M. (2012). Use of research-based instructional strategies in introductory physics: Where do faculty leave the innovation-decision process? *Physical Review Special Topics - Physics Education Research, 8,* 020104.

Herman, G. L., West, M., & Hahn, L. (2015). Coordinating college-wide instructional change through faculty communities. *International Mechanical Engineering Congress & Exposition*, 3-19.

Holgado-Tello, F., Chacón-Moscoso, S., Barbero–García, Is., & Vila, E. (2010). Polychoric versus Pearson correlations in exploratory and confirmatory factor analysis with ordinal variables. *Quality and Quantity, 44*, 153-166.

Johnson, E. (2019). Undergraduate mathematics instruction: Not as bad as you'd think? In *Levers for change: An assessment of progress on changing STEM instruction*, *6*(23), 100–120.

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2022). *semTools*: Useful tools for structural equation modeling. R package version 0.5-6. Retrieved from https://CRAN.R-project.org/package=semTools

Justice, N. (2020). Preparing graduate students to teach statistics: A review of research and ten practical recommendations. *Journal of Statistics Education, 28*, 334-343.

Justice, N., Zieffler, A., & Garfield, J. (2017). Statistics graduate teaching assistants' beliefs, practices and preparation for teaching introductory statistics. *Statistics Education Research Journal*, *16*(1), 294–319.

Katz, L. G. (1972). Developmental stages of preschool teachers. *The Elementary School Journal*, *73*(1), 50-54.

Keller, H. (2016). Psychological autonomy and hierarchical relatedness as organizers of developmental pathways. *Philosophical Transactions of the Royal Society B: Biological Sciences, 371*, 20150070.

Keller, M. M., Chang, M.-L., Becker, E. S., Goetz, T., & Frenzel, A. C. (2014). Teachers' emotional experiences and exhaustion as predictors of emotional labor in the classroom: An experience sampling study. *Frontiers in Psychology, 5*.

Kremer, T., Mamede, S., van den Broek, W. W., Schmidt, G. H., Nunes, P. M., & Martins, A. M. (2019). Influence of negative emotions on residents' learning of scientific information: An experimental study. *Perspectives on Medical Education, 8*, 209-215.

Kudryashova, A., & Rybushkina, S. (2016). Teacher's roles to facilitate active learning. *Mediterranean Journal of Social Sciences, 7*(1), 460-466.

Kulas, J., & Stachowski, A. (2013). Respondent rationale for neither agreeing nor disagreeing: Person and item contributors to middle category endorsement intent on likert personality indicators. *Journal of Research in Personality, 47*, 254–262.

Kulas, J., Stachowski, A. A., & Haynes, B. A. (2008). Middle response functioning in likert-responses to personality items. *Journal of Business and Psychology, 22*, 251-259.

Lee, J. W., Jones, P. S., Mineyama, Y., & Zhang, X. E. (2002). Cultural differences in responses to a likert scale. *Research in Nursing and Health, 25*, 295-306.

Leung, S. O. (2011). A comparison of psychometric properties and normality in 4, 5, 6, and 11-point likert scales. *Journal of Social Service Research, 37*, 412-421.

Little, T. (2013). *Longitudinal structural equation modeling.* New York, NY: Guilford Press.

Lyle, A. M., & Peurach, D. J. (2022). Changing notions of teacher autonomy: The intersection of teacher autonomy and instructional improvement in the US. *Research in Education*, 1-23.

Martens, R., Gulikers, J., & Bastiaens, T. (2004), The impact of intrinsic motivation on e-learning in authentic computer tasks. *Journal of Computer Assisted Learning, 20*, 368-376.

McAuley, E., Duncan, T., & Tammen, V. (1989). Psychometric properties of the intrinsic motivation inventory in a competitive sport setting: A confirmatory factor analysis. *Research Quarterly for Exercise Sport, 1*, 48-58.

McDonald R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Lawrence Erlbaum

Menold, J., Jablokow, K., Purzer, S., Ferguson, D., & Ohland, M. (2015). Using an instrument blueprint to support the rigorous development of new surveys and assessments in engineering education. *ASEE Annual Conference and Exposition*, 122.

Mezirow, J. (1991). *Transformative dimensions of adult learning*. Jossey-Bass.

Mezirow, J. (1997). Transformative learning: Theory to practice. In *New Directions for Adult and Continuing Education* (Vol. 74). Jossey-Bass.

Mezirow, J. (2003). Transformative learning as discourse. *Journal of Transformative Education, 1*, 58-63. doi:10.1177/1541344603252172

Miles, M. B., Huberman, A. M., & Saldana, J. (2014). *Qualitative data analysis: A methods sourcebook.* SAGE.

Mongillo, M. B. (2016). Preparing graduate students to teach math: Engaging with activities and viewing teaching models. *Networks: An Online Journal for Teacher Research, 18*(2), 1-19.

Nadler, J., Weston, R., & Voyles, E. (2015l). Stuck in the middle: The use and interpretation of mid-points in items on questionnaires. *The Journal of General Psychology, 142*, 71-89.

Nguyen, K. A., Borrego, M., Finelli, C. J., DeMonbrun, M., Crockett, C., Tharayil, S., . . . Rosenberg, R. S. (2021). Instructor strategies to aid implementation of active learning: A systematic literature review. *International Journal of STEM Education, 8*, 1-18.

Nias, J. (1996). Thinking about feeling: The emotions in teaching. *Cambridge Journal of Education, 26*, 293-306.

Norman, G. (2010). Likert scales, levels of measurement and the “laws” of statistics. *Advances in Health Sciences Education: Theory and Practice, 15*(5), 625-632.

Núñez, J. L., & León, J. (2019). Determinants of classroom engagement: A prospective test based on self-determination theory. *Teachers and Teaching, 25*, 147-159.

Obunadike, J. (2013). Development and validation of teacher motivation assessment scale for quality assurance in universities in Anambra state. *Journal of Educational and Social Research*, 95-104.

Ostinelli, G. (2016). The role of motivation and understanding in the change of teaching practices. *Journal of Inquiry and Action in Education, 7*, 111-128.

Park, J., Chung, S., An, H., Park, S., Lee, C., Kim, S. Y., . . . Kim, K.-S. (2012). A structural model of stress, motivation, and academic performance in medical students. *Psychiatry Investigation, 9*, 143-149.

Peterson, C. H., Peterson, N. A., & Powell, K. G. (2017). Cognitive interviewing for item development: Validity evidence based on content and response processes. *Measurement and Evaluation in Counseling and Development, 50*, 217-223. doi:10.1080/07481756.2017.1339564

Pfund, C., Miller, S., Brenner, K., Bruns, P., Chang, E.-M. D., . . . Handelsman, J. (2009). Professional development. Summer institute to improve university science teaching. *Science, 324*(5926), 470-471.

Piedmont, R. L. (2014). Inter-item correlations. In A. Michalos, *Encyclopedia of Quality of Life and Well-Being Research.* Springer.

Pinder, W. C. C. (2011). *Work motivation in organizational behavior* (2nd ed.). New York: Psychology Press.

Plutchik, R. (2001). The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. *American Scientist, 89*(4), 374-380.

Preston, C., & Colman, A. (2000). Optimal number of response categories in rating scales: Reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica, 104*, 1-15.

Prince, M. (2013). Does active learning work? A review of the research. *Journal of Engineering Education, 93*, 223-231.

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Reeve, J. (2002). Self-determination theory applied to educational settings. In E. L. Deci, & R. M. Ryan (Eds.), *Handbook of Self-Determination Research* (pp. 2–183). University of Rochester Press.

Ren, L., Green, J., & Smith, W. (2016). Using the Fennema-Sherman Mathematics Attitude Scales with lower-primary teachers. *Mathematics Education Research Journal, 28,* 303-326.

Rosseel, Y. (2012). *lavaan*: An R package for structural equation modeling. *Journal of Statistical Software, 48*(2), 1-36. http://www.jstatsoft.org/v48/i02/

Rossman, A. J., & Chance, B. L. (2021). *Rossman/Chance applet collection*. Retrieved from http://www.rossmanchance.com/applets/

Ryan, R. M. (1993). Agency and organization: intrinsic motivation, autonomy, and the self in psychological development. *Nebraska Symposium on Motivation, 40*, 1-56.

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology, 25,* 54-67.

Sabbag, A., Garfield, J., & Zieffler, A. (2018). Assessing statistical literacy and statistical reasoning: The reali instrument. *Statistics Education Research Journal*, *17*, 141-160.

Schellenbach-Zell, J., & Gräsel, C. (2011). Teacher motivation for participating in school innovations – supporting factors. *Journal for Educational Research Online, 2*.

Speer, N., Gutmann, T., & Murphy, T. J. (2005). Mathematics teaching assistant preparation and development. *College Teaching*, *53*, 75-80.

Stacy, A. M. (2000). The graduate student in the dual role of student and teacher. In *Graduate Education in the Chemical Sciences*, *Issues for the 21st Century: Report of a Workshop* (pp. 73–90). National Academy Press.

Suárez-Álvarez, J., Pedrosa, I., Lozano, L. M., García-Cueto, E., Cuesta, M., & Muñiz, J. (2018). Using reversed items in likert scales: A questionable practice. *Psicothema*, *30*(2), 149–158.

Sullivan, G. M., & Artino, A. R. (2013). Analyzing and interpreting data from likert-type scales. *Journal of Graduate Medical Education, 5*(4), 541-542.

Sutton, R. (2005). Teachers' emotions and classroom effectiveness: Implications from recent research. *The Clearing House, 78*, 229-234.

Teo, T., & Fan, X. (2013). Coefficient alpha and beyond: Issues and alternatives for educational research. *Asia-Pacific Education Researcher, 22*, 209-213.

Tishkovskaya, S., & Lancaster, A. G. (2012). Statistical education in the 21st century: A review of challenges, teaching innovations and strategies for reform. *Journal of Statistics Education, 20*(2), 1-56.

Tyng, C. M., Amin, H. U., Saad, M. N., & Malik, A. S. (2017). The influences of emotion on learning and memory. *Frontiers in Psychology, 8*, 1-22.

Vallerand, R. J. (1997). Toward A hierarchical model of intrinsic and extrinsic motivation. In M. P. Zanna (Ed.). *Advances In Experimental Psychology* (pp. 271-360), Academic Press.

Vickrey, T., Rosploch, K., Rahmanian, R., Pilarz, M., & Stains, M. (2017). Research-based implementation of peer instruction: A literature review. *CBE—Life Sciences Education, 14*(1), 1-11.

Von Treuer, K., & Reynolds, N. (2017). A competency model of psychology practice: Articulating complex skills and practices. *Frontiers in Education, 2,* 1-7.

Weihs, C., & Ickstadt, K. (2018). Data science: the impact of statistics. *International Journal of Data Science and Analytics, 6,* 189-194.

Wickham H (2016). *ggplot2: Elegant graphics for data analysis*. New York. Springer-Verlag. Retrieved from https://ggplot2.tidyverse.org.

Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology*, *25*(1), 68-81.

Wood, B. L., Mocko, M., Everson, M., Horton, N. J., & Velleman, P. (2018). Updated guidelines, updated curriculum: The GAISE college report and introductory statistics for the modern student. *CHANCE, 31*, 53–59. doi:10.1080/09332480.2018.1467642

Zembylas, M. (2005). Beyond teacher cognition and teacher beliefs: The value of the ethnography of emotions in teaching. *International Journal of Qualitative Studies in Education*, *18*, 465-487.

Zhu, J., & Thagard, P. (2002). Emotion and action. *Philosophical Psychology, 15,* 19–36.

Zieffler, A., Park, J., Garfield, J., delMas, R., & Bjornsdottir, A. (2012). The statistics teaching inventory: A survey on statistics teachers' classroom practices and beliefs. *Journal of Statistics Education*, *20*, 1-29.

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of Self-Regulation* (pp. 13-39). San Diego: Academic Press. doi:https://doi.org/10.1016/B978-012109890-2/50031-7

APPENDICES

APPENDIX A

PROGRAM SUPPORT FROM CHAPTER 1

Orientation Week

Building Community – GSIs first discussed their goals as educators and how they might achieve these goals within their classrooms. Then the discussion focused on the importance of building a classroom community and sharing various strategies GSIs can use to build a community in their classroom.
The First Day of Class – GSIs discussed ways to get their class off to a good start and how to establish their course expectations and standards. GSIs discussed their ideal classroom environment and identified strategies to facilitate those types of classroom environments.
Engaging Students in the Classroom – GSIs learned and demonstrated instructional techniques and strategies to actively engage their students. Working with faculty, GSIs first watched a short mathematics lesson on factoring which employed a lecture-based approach of instruction. After the video, the GSIs prepared and presented a short demonstration that conveyed the same material in a way that involved active learning strategies.
Teaching Demonstrations – GSIs prepared and presented a small portion of a lesson they would teach during the first week of class. Facilitators and other GSIs received feedback and discussed the teaching demonstrations.
Panel Discussion – Experienced GSIs provided advice to new GSIs and answered any lingering questions from new GSIs.
Fall Semester
Facilitating Group Work – GSIs discussed strategies they can use to facilitate group work in their classrooms. Additionally, GSIs reflected on common challenges that occur when students work in groups, and they identified ways to address these challenges in their classrooms.
Structuring a Lesson – GSIs explored different ways to launch and close a lesson, and they discussed the importance of these components in a lesson. Then, GSIs practiced creating and delivering a launch and closure for different mathematics/statistics topics.
Active Learning Teaching Strategies – GSIs learned about and practiced various active learning teaching strategies.
Formative Assessment – GSIs explored different ways to informally assess student understanding. They examined sample student work to holistically assess what knowledge students were and were not yet demonstrating and strategized how to use this information to adapt their instruction.
Approachability – Timed after GSIs gave their first exam, GSIs discussed how they can encourage their students to attend office hours and receive additional help. GSIs were then given the opportunity to role-play and practice their strategies with others.
Classroom Observations – GSIs first identified a teaching strategy (such as group work) they were interested in learning about and then observed a pre-selected faculty member enact that strategy in their classroom. Lastly, GSIs reflected on what they observed.
Formal and Informal Evaluations – GSIs brainstormed ideas for informal evaluation questions to obtain feedback from their students and reviewed the formal course evaluation form that students complete at the end of the semester in their courses.
Using Technology in the Classroom – GSIs viewed and reflected on strategies for using technology to clarify and promote students' mathematical/statistical understanding outside of the classroom.
Responding to Student Contributions – GSIs learned and practiced strategies for asking questions during instruction and responding to student responses. For example, GSIs discussed when (and how) instructors can answer questions with questions to encourage student learning.
The Journey to Academia – A panel of mathematics and statistics instructors shared their different paths to becoming teachers and provided advice for those interested in teaching as a career.

First Semester Reflections – GSI reflected on their first semester teaching and what went well and what they hope to improve for the spring semester.
Spring Semester
Reflecting on Course Evaluations – GSIs received their student course evaluations from the fall semester and discussed how to constructively use those student comments to improve teaching. GSIs also discussed how personal biases can affect how others perceive instructors, students and learning.
Curriculum Vitae and Resumes – GSIs brought their current CVs and resumes and received feedback from faculty and experts in industry.
Practical Teaching Strategies – An experienced instructor shared practical teaching tips for enhancing classroom instruction and student engagement. GSIs also participated in a discussion regarding what happens when instructors try something new, and it doesn't work in the way they hoped.
Teaching Inclusively – GSIs reflected on what it means to teach inclusively. They discussed strategies that can be used to help students (perhaps even themselves) feel like valued, contributing members of their classroom communities.

APPENDIX B

INTERVIEW PROTOCOL FROM CHAPTER 1

Our observation protocol included initial expansive questions that helped facilitate flexible conversation around GSIs' knowledge about, emotions towards, and use of active learning.

Knowledge: Information an individual holds about what active learning is

- What is your definition of active learning?
- What different types of teaching techniques would you consider active learning?
- What is your role during these types of techniques?

Emotions: Mental states brought on by active learning

- Do you find active learning necessary?
- Do you find active learning useful?
- Do you find active learning beneficial?

Use: How active learning is implemented in the classroom

- What do you consider when planning active learning?
- What do you pay attention to when using active learning?
- Do you reflect on your use of active learning?

APPENDIX C

SURVEYS GIVEN FROM CHAPTER 1

2017 Pre-Survey

Name:

1. How do you learn math/statistics best?
2. How do you think your students will learn math/statistics best?
3. How do you describe your role as a teacher?
4. How do you know when your students understand?

2017 Post-Survey

Name:

1. How have you used active learning this semester in your class?
2. Which types of active learning techniques do you plan to use next semester?

2018 Survey

1. How you define active learning? Has your definition changed since last semester?

2. How do you feel about active learning and its place in the classroom? What is your reaction when challenged to implement a different active learning technique? During lesson prep, how do the different types of active learning techniques planned for the class alter your outlook/mentality on that day’s lecture?

3. What new active learning techniques have you learned this semester (if any)? What current active learning strategies have you been using in your classroom? What active learning strategies would you like to use as the semester progresses?

4. Describe how effective you perceive yourself in implementing different types of active learning techniques. Are there any limitations that are holding you back from implementing different strategies? What actions are you taking to make different techniques effective?

5. Explain why you use these active learning strategies. How do you think these active learning strategies benefit students, if at all?

2019 Survey

Please answer the following questions below to the best of your ability.

1. Please list the following courses you have taught at Montana State University, including which semester(s) and year(s) you've taught each.
2. When you first discovered active learning (here at MSU or another facility), what challenge(s) did you face when trying to implement such techniques?
3. During your latest semester teaching, what challenge(s) did you face when trying to implement active learning techniques?

APPENDIX D

OBSERVATION PROTOCOL FROM CHAPTER 1

<u>Name of observer</u>
<u>Name of observation</u>

Class:
Topic:
Length of Class:

Summary:

Detailed summary of both the frequency and use of active learning for the duration of the lesson.

Active Learning Categories:

List of active learning techniques used, categorized in sheet given. Observations may include more than one categorized active learning technique and should be listed as such. Rough time stamps for identified active learning techniques should be included.

Student Engagement:

Summary of student engagement throughout for the duration of the lesson.

Facilitation:

Facilitation is a term used to describe a possible role of the teacher. Facilitation is providing the necessary resources, information and support in order for learners to complete a task, rather than teaching.

Summary of active learning instructional use.

APPENDIX E

ITEM BLUEPRINT FROM CHAPTER 4

Response Type	Likert Scale: 1 – 6 Strongly agree Agree Slightly agree Slightly disagree Disagree Strongly disagree
Instruments Reviewed when Creating Items	The Student Motivation Scale (SMS) (Martin, 2001) Developing an Instrument to Measure Motivation, Learning Strategies and Conceptual Change On the Assessment of Situational Intrinsic and Extrinsic Motivation: The Situational Motivation Scale (SIMS) (Guay et al., 2000) The perceived autonomy support scale for exercise settings (PASSES) (Hagger et al., 2007) The Statistics Teaching Inventory: A Survey on Statistics Teachers' Classroom Practices and Beliefs (Zieffler et al., 2012) Assessing Statistical Literacy and Statistical Reasoning: The REALI Instrument (Sabbag et al., 2018) Using the Fennema-Sherman Mathematics Attitude Scales with lower-primary teachers (Ren et al., 2016)
Type of Motivation	**Intrinsic motivation -** Performing an activity for itself, in order to experience pleasure and satisfaction inherent in the activity. **Identified regulation -** Non-intrinsically motivated behavior has been completely internalized, there is no internal resistance, and there is a willingness to do the behavior because it is important and valuable to the self, albeit it does not have to be enjoyable. **External regulation -** Behavior that has not been internalized. A willingness to do a behavior to obtain something tangible that satisfies an eternal demand or avoids punishment. **Amotivation -** The absence of both extrinsic and intrinsic motivation. One engaged in a behavior because they are " going through the motions." (Ryan & Deci, 2000)
Active Learning Technique	**Group work -** When using group work for active learning, students communicate, share ideas, and think critically about the topic(s) with their group members. This includes think-pair-share, group presentations, or other small group work activities that have students engage with their group members and the topic(s) as described above. **Using Technology -** Technology refers to technological tools that assist in the analysis of data, communication of ideas, and development of student understanding. When using technology for active learning, students acquire information and discover statistical ideas through their interaction with the technological tool. This may include having students work with Tableau, CODAP, R, Tinkerplots, applets, etc. to discover concepts. This does not include using passive technology, such as displaying a PowerPoint or using a calculator for calculations **Real data** – Data that is not fake or simulated. Using real data may include collecting data from students during class or preparing real world data to integrate into a lesson that focuses on the data' s context and purpose. Collecting data may involve the administration of an in-class survey or an out-of-class survey to obtain information from students. **Large-group discussions** – Conversation about the topic(s) at the class level. Large-group discussions are designed to help students think about and

APPENDIX F

PHASE 1 EXPERT OPINION SURVEY FROM CHAPTER 4

Note: Participation in this study is voluntary, and you are not obligated to submit the survey at any time.
The purpose of this research study is to develop an instrument to measure graduate student instructors' (GSIs') motivation to use active learning teaching techniques. The instrument's target population is GSIs who teach an introductory statistics course or a recitation section.
Within the context of this study, we are currently defining active learning and motivation in the following way:

Active learning refers to "classroom practices that engage students in activities, such as reading, writing, discussion, or problem solving, that promote higher-order thinking" (CBMS, 2016, 1).

Motivation is a multi-dimensional construct characterizing why individuals choose to use (or not use) active learning teaching techniques. Motivation is the "why" of behavior (Deci & Ryan, 1985).

There are many different types of active learning techniques GSIs may use when teaching introductory statistics. The purpose of this survey is to gather experts' opinions about which active learning techniques are most valuable to include on an instrument measuring GSIs' motivation. Your opinions will help us identify which techniques to address when creating research instrument items.

1. At your institution, what are the teaching roles and responsibilities of graduate student instructors (GSIs) who teach statistics? Please also state whether GSIs at your institution are sole instructors of any statistics courses and, if so, which ones. If you do not have GSIs at your institution, please type, "We do not have GSIs at our institution."

2. Based on the GAISE Guidelines and other literature on active learning in statistics classrooms, we have selected four activities for you to review. These are:

Group work – Method of instruction that gets students to work together in groups of two or more. Group work involves strategies that allow students to communicate with peers, share their ideas, and think critically about the topic(s). This may include think-pair-share, group presentations, or other small group work activities that have the characteristics described above.

Technology – Technological tools that assist in the communication, development, and exchange of knowledge. Using technology is about designing a lesson that allows students to acquire information through discovering material for themselves. This may include having students work with Tableau, CODAP, R, etc. to discover information. This does not include passive technology, such as displaying a power point.

Real data – Data that is not fake or simulated. Using real data may include collecting data from students during class or preparing real world data to integrate into a lesson that focuses on the data's context and purpose. Collecting data may involve the administration of an in-class survey or an out-of-class survey to obtain information from students.

Large-group discussions – Conversation about the topic(s) at the class level. Large-group discussions are designed to help students think about and express their ideas with others in the class. During discussion, instructors prepare open-ended questions and move the discussion forward by having students elaborate on their thinking through providing explanations, evidence, or clarifications, and inviting others to react and respond by providing similar and/or alternative viewpoints.

a. These definitions will be provided to GSIs when filling out the research instrument. Please review these definitions and answer the following questions:

- Do you agree with each definition? If not, please explain.

- Do you find these definitions specific enough to clearly describe these activities to a general graduate student teaching audience? If not, please explain.

b. If applicable, please use the space below to refine the definitions and address any concerns you noticed.

3. Please list any other active learning techniques that you would like us to consider having on the research instrument we are developing. Please include a working definition and description of each active learning teaching technique you list.

4. Please rank the following active learning techniques (including your own listings) in order of which techniques you would like to be included on an instrument measuring GSIs' motivation to engage in active learning techniques. Assign a value of 1 to the active learning technique in which you have the largest interest, then continue numbering in order of preference until you have reached the total number of active learning techniques.

5. Would you be willing to serve as an expert reviewer of drafted instrument items? As an expert reviewer, you would be asked to assess the validity of the items in relation to the chosen active learning techniques, as well as identify potential concerns or issues with each item's wording. We expect to complete a draft of these items by (###) and hope to have the expert review of the items completed by the end of Spring 2020 semester. If willing, you will be sent an email with more information about the items and the review process. Thank you for your support in advancing this research.

6. (Optional) If willing, please list the names and contact information of others you recommend contacting to complete this survey and/or review a draft of research instrument items.

APPENDIX G

PHASE 2 EXPERT ITEM REVIEW SURVEY FROM CHAPTER 4

Consent for Participation & Purpose of Study

Consent

Participation in this study is voluntary. You are free to stop participating and may withdraw your consent at any time. You are not obligated to submit the survey, and you may skip any questions in the survey you want. There are no foreseen risks or benefits to you as a participant. We will not identify you by name in any reports using information obtained in the survey, and your confidentiality as a participant in this study will remain secure.

Contact Information

If you have any questions about the survey or this research project, you may contact me (elijah.meyer@montana.edu), Jennifer Green (jg@msu.edu), or Stacey Hancock (stacey.hancock@montana.edu). If you have additional questions about the rights of human subjects, you may contact the Chair of the Institutional Review Board, Mark Quinn (mquinn@montana.edu).

Study Description and Purpose

The purpose of this study is to develop an instrument to measure graduate student instructors' (GSIs') motivation to use active learning teaching techniques. We define GSIs as graduate students who are the sole or lead instructor of a statistics course or lead a recitation section. This instrument will measure four different types of motivation to use active learning on GSIs when teaching statistics. These motivations include intrinsic, integrated regulation, external regulation, and amotivation. These four types of motivation have been selected because they span our framework of motivation and are linked to a variety of behavioral outcomes in other areas of research.

The purpose of this survey is to gather experts' feedback on the written items. Your feedback will help identify any opportunities to improve wording and help provide initial content evidence suggesting that these items are measuring their intended type of motivation.

Part 1: Item Review

Directions

Based on feedback from the last survey, two different types of active learning strategies were selected to be included on our instrument: group work and use of technology. You are asked to please review items for one of the two selected strategies: *Group work.* I have provided the written description of the selected strategy that GSIs will read before responding to the corresponding items of motivation below.

We ask that you provide feedback on three different areas on this instrument:

- Feedback on the active learning strategy's definition
- Feedback on the drafted items
- Feedback on the motivational constructs' relationship with the drafted items

Questions targeting these three different areas are provided to help guide the review process. These questions can be found in the *Feedback Questions* sections throughout the document. When providing feedback, please note that this **instrument is intended for GSIs who use active learning while teaching statistics.** Please provide feedback in the appropriate feedback sections below, or through inserted comments within the document.

Additionally, we ask you to please review the background questions intended to collect additional data on GSIs. This can be found at the end of the survey in *Part 2*.

Active Learning Strategy 1 – Group work

Group work refers to a method of instruction that gets students to work together in groups of two or more.

When using group work for active learning, students communicate, share ideas, and think critically about the topic(s) with their group members. This includes think-pair-share, group presentations, or other small group work activities that have students engage with their group members and the topic(s) as described above.

Feedback Questions: Definition

- Do you agree with the working definition of group work? If not, please explain.
- Do you find this definition specific enough to clearly describe these activities to a general graduate student teaching audience? If not, please explain.

Definition Feedback:

Construct 1: Intrinsic Motivation

Intrinsic Motivation – Performing an activity for oneself, to experience pleasure and satisfaction inherent in the activity. Example: A person plays a game of basketball because of the sheer joy they experience while playing it.

Items

Directions: Read each item carefully. Using the scale below, please select the number that best reflects the extent to which you agree or disagree with the following statements about your use of technology in the classroom when teaching statistics.
Answer each item according to the following scale: 5 – Strongly Agree; 4 – Agree; 3 – Neutral; 2 – Disagree; 1 – Strongly Disagree

a. I find using group work satisfying when teaching statistics
b. I find it enjoyable to use group work when teaching statistics

c. I find it interesting to use group work when teaching statistics
d. I'm committed to regularly using group work when teaching statistics
e. I get excited when using group work to teach statistics
f. My favorite statistics lessons to teach are ones that involve group work
g. I feel proud about lessons where I have used group work
h. I feel pleased about lessons where I have used group work

Feedback Questions: Wording

- Are there concerns about the wording of the item (e.g., double-barreled, idioms, jargon, etc.)?
- Are there concerns about how items are phrased (e.g., leading items, items that may be misinterpreted, items that may not be appropriate for GSIs)?

Wording Feedback

Feedback Questions: Construct

- Do you believe each item reflect qualities of intrinsic motivation?
- Do you believe there are missing items?
- Are aspects of intrinsic motivation's definition mis- or underrepresented within the set of items?

Construct Feedback

Construct 2: Integrated Regulation

Integrated Regulation – A willingness to engage in a behavior because it is important and valuable to oneself. Example: A person attends school because they believe that the act aligns with their personal belief system.

Items

a. Using group work is necessary when teaching statistics
b. Using group work is consistent with my goals as an instructor when teaching statistics
c. Using group work is consistent with my values as an instructor when teaching statistics
d. Using group work is essential to my identity as an instructor when teaching statistics
e. Using group work is important for me as an instructor when teaching statistics
f. I incorporate group work to align my teaching with disciplinary "best" practices
g. Using group work makes me a better instructor when teaching statistics

- Feedback Questions: Wording Are there concerns about the wording of the item (e.g., double-barreled, idioms, jargon, etc.)?
- Are there concerns about how items are phrased (e.g., leading items, items that may be misinterpreted, items that may not be appropriate for GSIs)?

Wording Feedback:

Feedback Questions: Construct

- Do you believe each item reflect qualities of integrated regulation?
- Do you believe there are missing items?

- Are aspects of integrated regulation's definition mis- or underrepresented within the set of items?

Relationship Feedback:

Construct 3: External Regulation

External Regulation – Engaging in a behavior to satisfy an external demand, receive an external reward, or avoid a punishment. Example: A student studies hard to get a good grade and receive a reward from their parents, or to avoid punishment for receiving a poor grade.

Items

a. I use group work to teach statistics because it is recommended to me by my colleagues
b. I use group work to teach statistics because this is the way I am expected to teach statistics by my colleagues
c. I use group work to teach statistics because I would get in trouble by my supervisors if I didn't use it
d. I use group work to teach statistics because students would be upset with me if I didn't teach using it
e. I use group work to teach statistics because I think it will help me earn a teaching award
f. I use group work to teach statistics because I am praised by my colleagues for doing so
g. I use group work to teach statistics only when I am being observed by another instructor
h. I use group work to teach statistics so that my peers think I am a good instructor

Feedback Questions: Wording

- Are there concerns about the wording of the item (e.g., double-barreled, idioms, jargon, etc.)?
- Are there concerns about how items are phrased (e.g., leading items, items that may be misinterpreted, items that may not be appropriate for GSIs)?

Wording Feedback:

Feedback Questions: Construct

- Do you believe each item reflect qualities of external regulation?
- Do you believe there are missing items?
- Are aspects of external regulation's definition mis- or underrepresented within the set of items?

Relationship Feedback:

Construct 4: Amotivation Items

Amotivation – The absence of intention or clear motives to engage in an activity. Amotivation is the lack of intrinsic motivation, integrated regulation, and external regulation. Example: An athlete claims to not value or see the point in training anymore for competition.

Items

a. Using group work to teach statistics does not improve the way statistics is taught
b. Using group work to teach statistics does not improve student learning

c. I question if I should use group work to teach statistics
d. Group work's role in teaching statistics is unclear to me
e. Using group work to teach statistics ends up using more class time than it is worth
f. I do not see what value using group work to teach statistics brings to my classroom
g. I do not see what value using group work to teach statistics bring to me as an instructor
h. I do not wish to continue incorporating group work in my classroom
i. I do not know why I use group work to teach statistics
j. I do not see myself using group work in the future while teaching
k. I do not understand why we use group work when teaching statistics

Feedback Questions: Wording

- Are there concerns about the wording of the item (e.g., double-barreled, idioms, jargon, etc.)?
- Are there concerns about how items are phrased (e.g., leading items, items that may be misinterpreted, items that may not be appropriate for GSIs)?

Wording Feedback:

Feedback Questions: Relationship

- Do you believe each item reflect qualities of amotivation?
- Do you believe there are missing items?
- Are aspects of amotivation's definition mis- or underrepresented within the set of items?

Relationship Feedback:

Additional Comments

If you have additional comments that you would like to make about the sections above, please do so here:

Part 2: Background Questions Review

Directions – Background Questions Review

Background questions will be given at the beginning of the instrument to collect background data on each GSI filling out the survey. We ask you to please review these background questions and consider the following:

Feedback Questions

- Are the background questions written clearly?
- Are the background questions appropriate for GSIs?
- Are GSIs unable to answer certain background questions?

- Are there any additional background questions you suggest we consider collecting data on?

Please write feedback in the *Background Questions Feedback* area at the end of the background questions, or through inserted comments within the section.

Background Questions

1. Are you currently a graduate student or within one year of graduation?
2. Please list the name of the university in which you are currently enrolled as a graduate student.
3. Please list the name of the department in which you are currently enrolled as a graduate student.

4. What degree type you are currently pursuing? Ex. Doctorate in Statistics

5. What is the subject you are getting your degree in? Ex. Statistics

6. How many years have you been a graduate student at your current institution? Round up (e.g., if you have completed 3.25 years, please enter 4).

7. Are you currently involved in teaching a statistics course, or leading a statistics recitation/discussion section?

8. Prior to becoming a graduate student at your current institution, did you have experience teaching as an instructor or recitation/discussion leader?

9. How many college terms you have taught a statistics course or led a statistics recitation/discussion section before being enrolled as a graduate student at your current institution? Ex. 2 Semesters; 3 Quarters

10. How many terms you have taught a statistics course or led a statistics recitation/discussion section while being enrolled as a graduate student at your current institution? If you are currently teaching a statistics course or leading a statistics recitation/discussion section this term, please include that in your count

11. List all statistics courses and statistics recitation/discussion sections you have taught and indicate how many times you have taught each course or section.

12. Do you engage in conversations about teaching with other peers? If so, how often and what about?

13. Describe the classroom that you currently teach in. In other words, describe the arrangement of desks/tables, the technology available, Etc.

14. Does your department offer a training program for new graduate students?

15. What kind of support does your department provide graduate student instructors?

16. Please read the following definition of using technology in your classroom:

Technology refers to technological tools that assist in the analysis of data, communication of ideas, and development of student understanding.
When using technology for active learning, students acquire information and discover statistical ideas through their interaction with the technological tool. This may include having students work with Tableau, CODAP, R, Tinkerplots, applets, etc. to discover concepts. This does not include using passive technology, such as displaying a PowerPoint or using a calculator for calculations.

Based on this definition, describe how you use technology when teaching statistics. If you do not use technology when teaching statistics, please write "N/A."

17. Please read the following definition of using group work in your classroom:

Group work refers to a method of instruction that gets students to work together in groups of two or more.
When using group work for active learning, students communicate, share ideas, and think critically about the topic(s) with their group members. This includes think-pair-share, group presentations, or other small group work activities that have students engage with their group members and the topic(s) as described above.

Based on this definition, describe how you use group work when teaching statistics.

Background Question Feedback

Additional Comments

If you have additional comments that you would like to make about the background questions above, please do so here:

APPENDIX H

PHASE 3 INTERVIEW PROTOCOL FROM CHAPTER 4

General Questions to Guide Cognitive Interview. Interview protocol consisting of different types of questions (probes) to incite conversation around items.

Type of Probe	Questions
Expansive	• What does this mean to you? • Tell me more about what this item means to you? • Please explain your thought process when answering this item. • Please explain how you answered this item. • Please explain what you believe this item is asking.
Anticipated	• Did you find anything confusing or unclear about this item? • How do you interpret the word *regularly?* • How do you define the term Collogues / Supervisors / People? • Did the lack of a neutral option noticeably affect your decision making? • How was the length of the survey? • Was there a point in the survey when you started to feel fatigued? • Was there a point in the survey did you started to "skim" or read faster than normal? • Is this question appropriate for you to answer? • How did the survey flow? Would you have liked to answer certain background questions before others? • Did you feel the survey was repetitive? • What questions / items were the hardest to answer? • How would you respond if a colleague asked you to fill out this instrument administered this to you during the semester? • *Technology Question*: What other types of technology, that you don't use, would you consider here? • *Group work:* What other types of group work, that you don't use, would you consider here?
Conditional	• Please explain why you marked # for this item; Condition - if inconsistent or extreme response is observed • Was this item / set of items difficult to answer; Condition – Observed long pauses or if respondent takes an extended period of time to answer an item / sets of items. • What would make this item more relatable to you?; Condition – If item is unable to be answered because it is deemed inappropriate • Please explain why you marked # for these sets of item; Condition - if pattern across items is observed. May indicate evidence of survey fatigue, or disinterest. • Was this item difficult to answer; Condition – Verbal or non-verbal clue signaling an abnormal response time or reaction to an item.
Things to look for during interview	• Change in verbal expressions • Contradictory statements among similar items • Refusal to answer items • Reluctant to answer items

APPENDIX I

R-CODE APPENDIX FROM CHAPTER 4

Full code for Phase 4 and Phase 5 is provided electronically at

https://github.com/ElijahMeyer3/Appendix-R-code.

APPENDIX J

PHASE 4 & 5 RESULTS

Phase 4: Summary Statistics of Intrinsic Motivation Items

	IM1	IM2	IM3	IM4	IM5	IM6	IM7	IM8
Min	3	4	3	3	3	3	3	3
Q1	5	5	4	4	4	4	4	4
Median	5	5	5	5	5	5	4	4
Mean	5.18	5.09	4.45	4.73	4.64	4.73	4.54	4.45
Q3	6	5.5	5	5	5	5	5	5
Max	6	6	6	6	6	6	6	6

Phase 4: Summary Statistics of Integrated Regulation Items

	IR1	IR 2	IR 3	IR 4	IR 5	IR 6	IR 7	IR 8	IR 9	IR 10	IR 11
Min	4	2	3	4	4	3	4	4	3	3	3
Q1	5	4	5	5	5	4	5	4	4	3.5	4
Median	6	5	5	5	5	5	5	5	5	5	5
Mean	5.45	4.54	5.00	5.18	5.36	4.91	5.09	4.63	4.73	4.54	4.82
Q3	6	5	5.5	6	6	6	5.5	5	5.5	5.5	6
Max	6	6	6	6	6	6	6	6	6	6	6

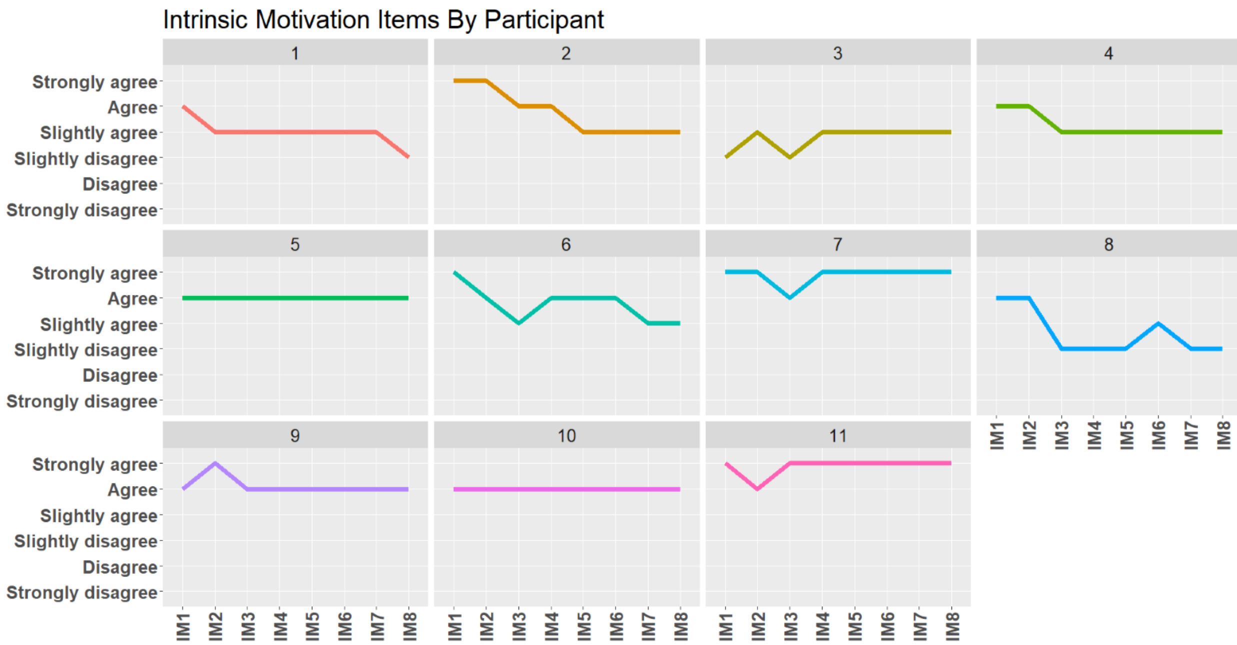
Phase 4: Summary Statistics of External Regulation Items

	ER1	ER2	ER3	ER4	ER5	ER6	ER7
Min	1	3	4	1	1	1	2
Q1	3.5	3.5	5	3	3	2	2
Med	5	5	5	4	4	3	3
Mean	4.54	4.36	5.27	4.09	3.45	2.91	2.91
Q3	6	5	6	5.5	4	3	3
Max	6	6	6	6	5	6	5

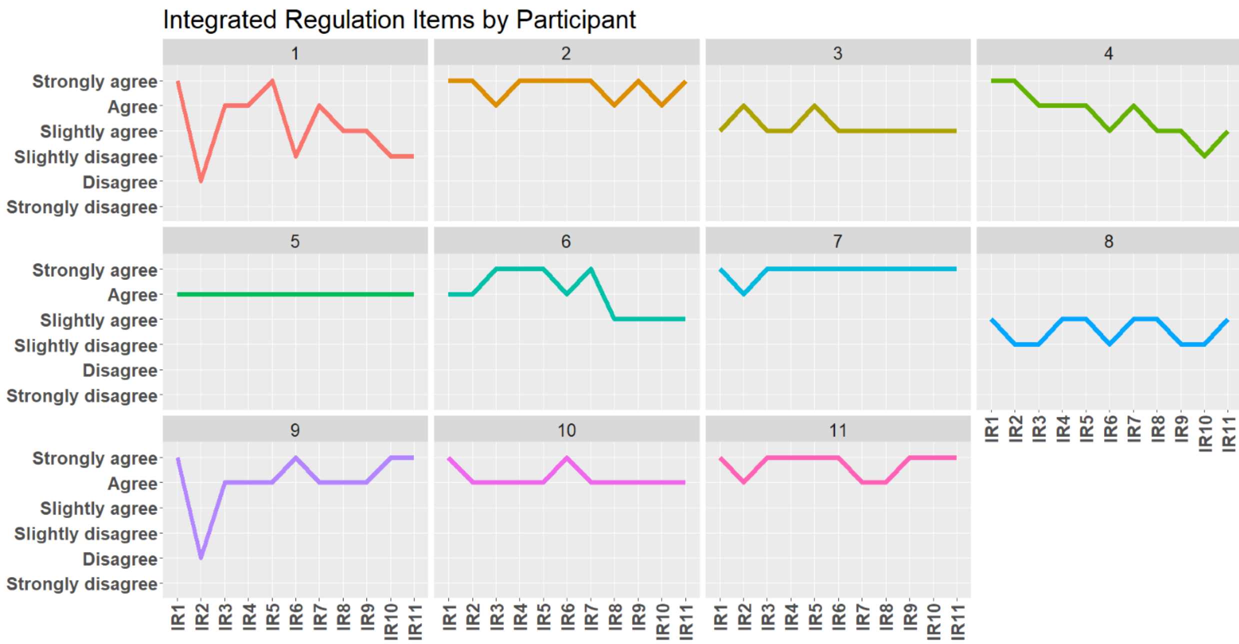
Phase 4: Summary Statistics of Amotivation Items

	AM1	AM2	AM3	AM4	AM5	AM6	AM7	AM8
Min	1	1	1	1	1	2	1	1
Q1	1	2	1	1.5	1	2	1	1
Med	1	2	2	2	2	3	2	2
Mean	1.45	2.55	2.36	2.45	2.09	3.18	2	1.91
Q3	2	3	3.5	3.5	3	4.5	2.5	2.5
Max	3	5	5	4	4	5	4	3

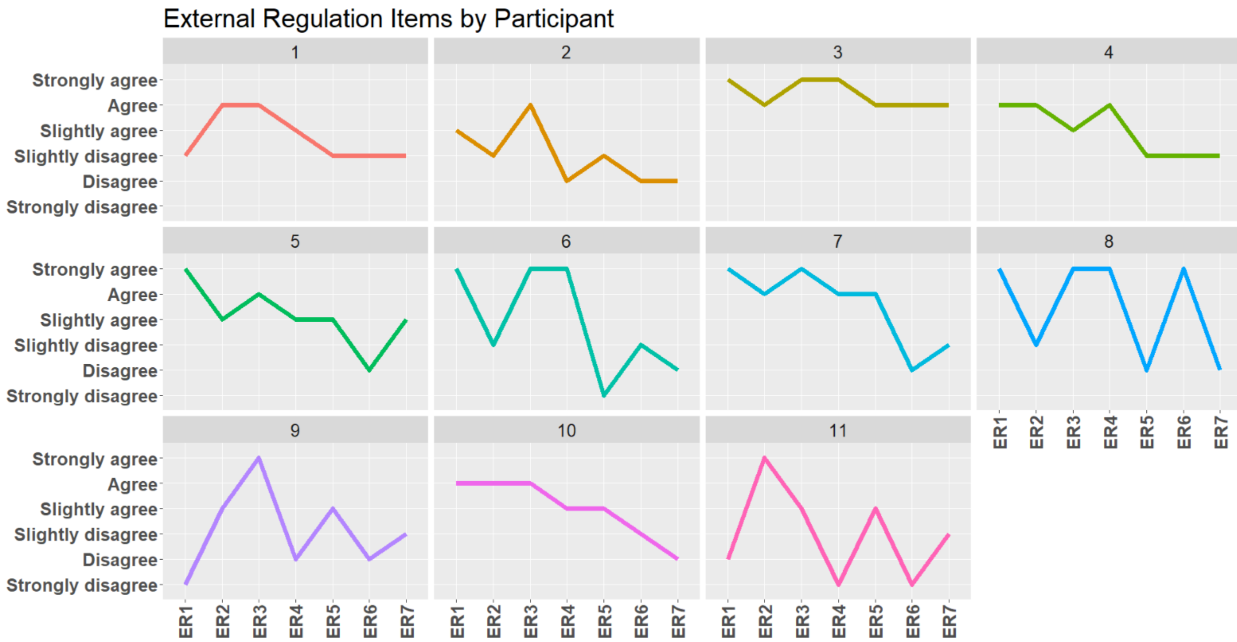
Phase 4: Line Graphs of Intrinsic Motivation Items



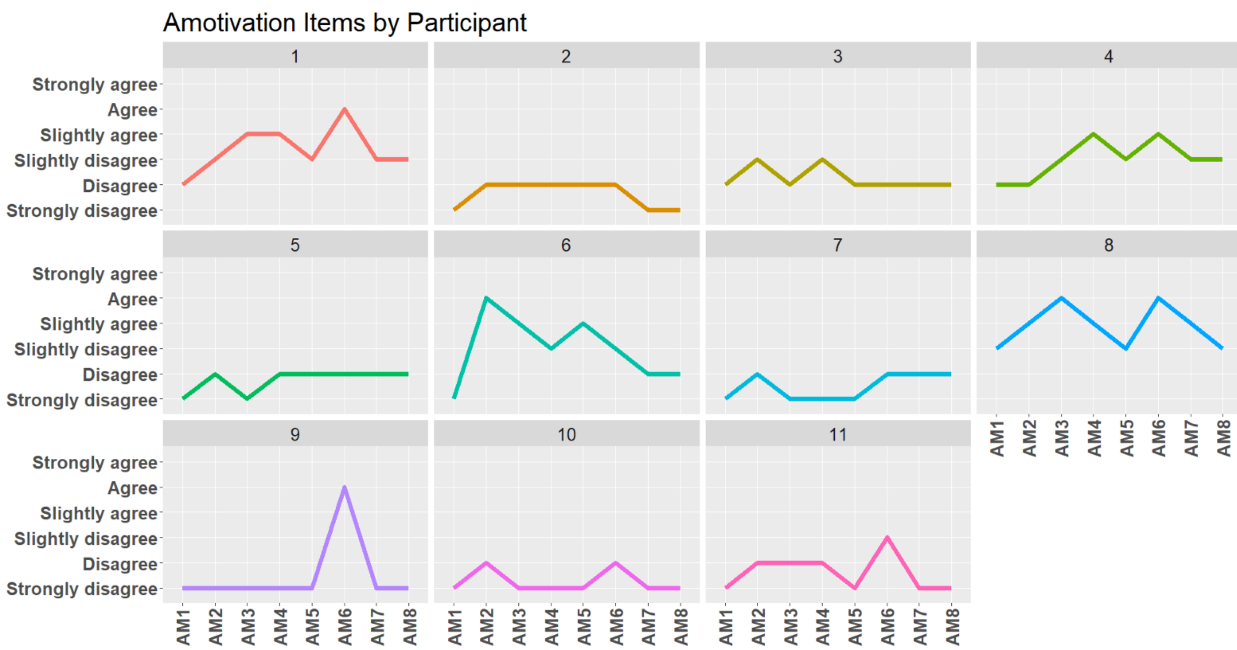
Phase 4: Line Graphs of Integrated Regulation Items



Phase 4: Line Graphs of External Regulation Items



Phase 4: Line Graphs of Amotivation Items



Phase 4: Bootstrapped 95% Confidence Intervals for Cronbach's Alpha Estimates

Motivational Construct	**Cronbach's Alpha Estimate**	**Lower Bound**	**Upper Bound**
Intrinsic Motivation	0.952	0.852	0.972
Integrated Regulation	0.929	0.773	0.975
External Regulation	0.569	0.018	0.878
Amotivation	0.929	0.805	0.971
Rewards	0.809	0.346	0.904
Pressure	0.826	0.623	0.928

Phase 4: Inter-item Correlations Table

	Intrinsic Motivation	**Integrated Regulation**	**Rewards**	**Pressure**	**Amotivation**
Intrinsic Motivation					
Integrated Regulation	0.61				
Rewards	-0.05	-0.10			
Pressure	-0.26	-0.26			
Amotivation	-0.55	-0.52	0.19	-0.07	

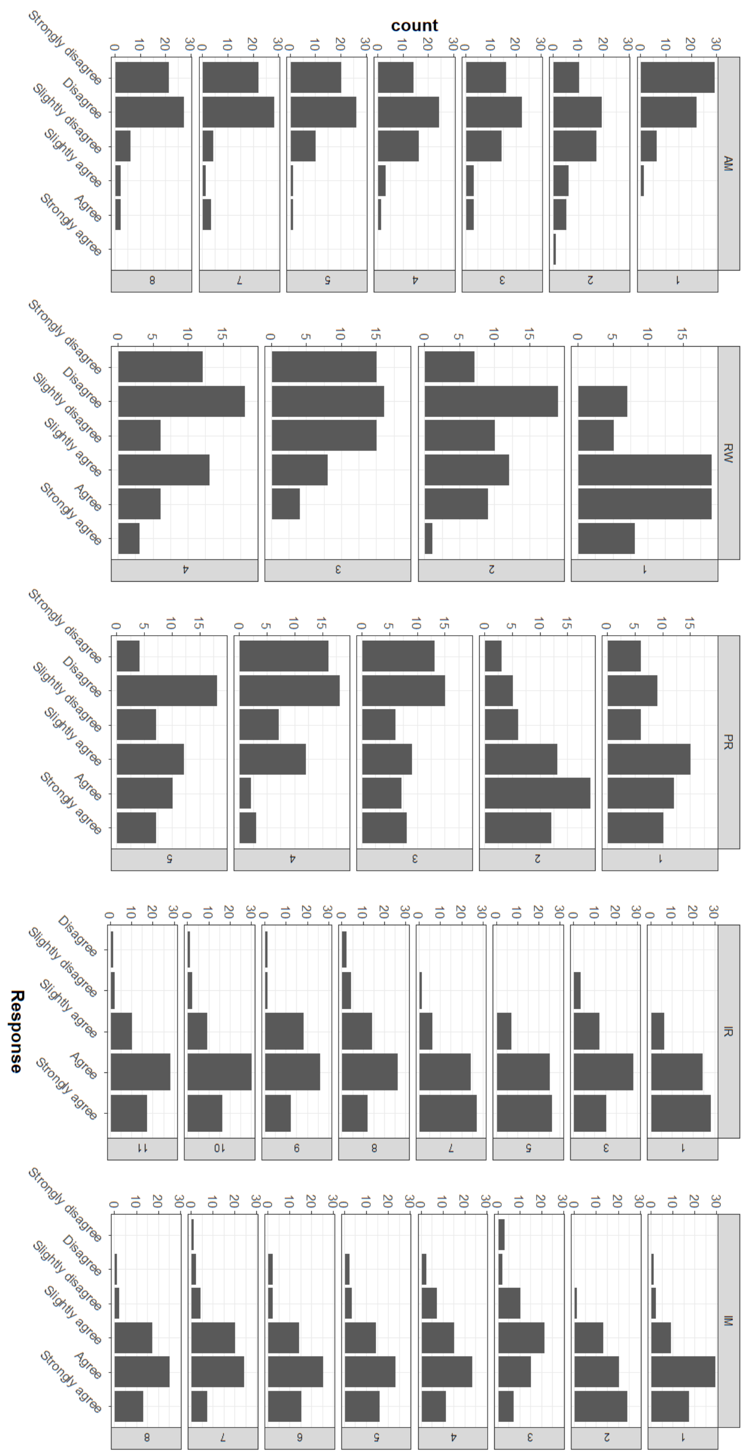
Phase 5: Maximum Likelihood Estimated 95% Confidence Intervals for Omega Estimates

Motivational Construct	**Coefficient Omega Estimate**	**Lower Bound**	**Upper Bound**
Intrinsic Motivation	0.925	0.894	0.956
Integrated Regulation	0.925	0.885	0.965
Pressure	0.931	0.906	0.956
Rewards	0.824	0.749	0.900
Amotivation	0.811	0.730	0.893

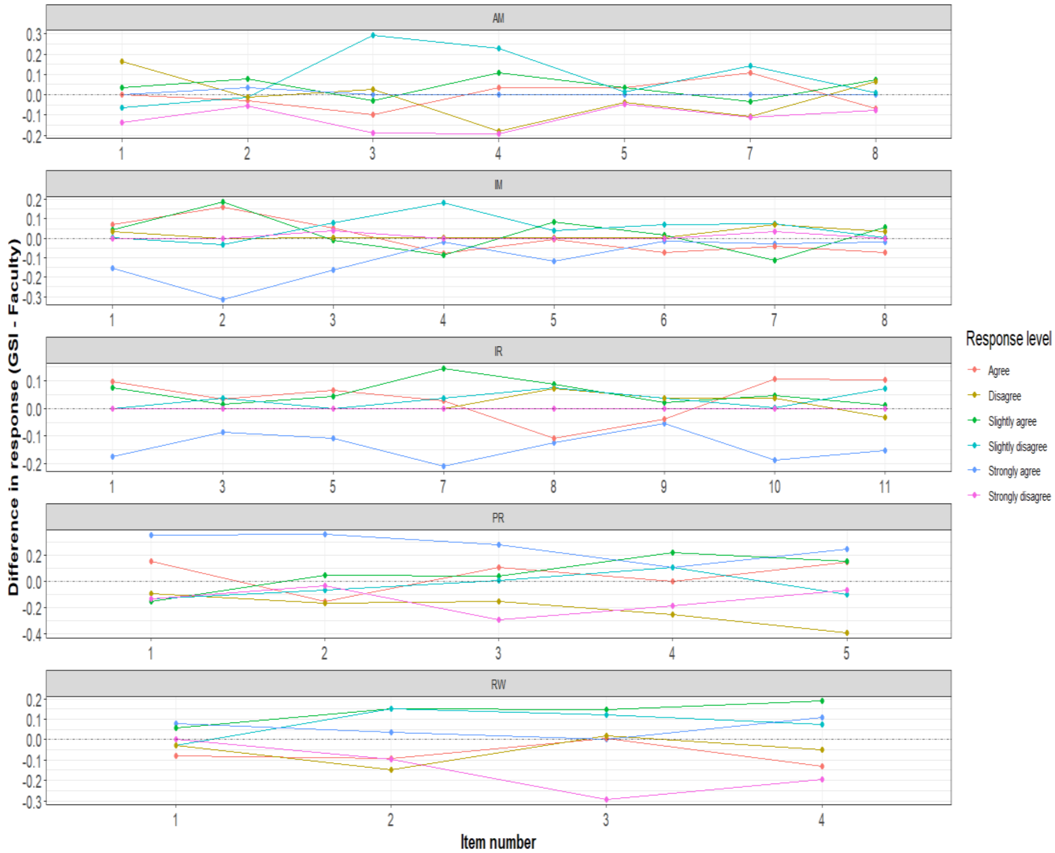
Phase 5: Fit Statistics from CFA Model

Chi-Square	Statistic: 866.237 p-value: < 0.001
Root Mean Square Error of Approximation	0.125

Histogram of Likert Scale Responses for Phase 5



Phase 5: Line Graph of Difference in Proportion of Responses by GSI vs non-GSI



Phase 5: Factor Correlations Table

	Intrinsic Motivation	**Integrated Regulation**	**Rewards**	**Pressure**	**Amotivation**
Intrinsic Motivation					
Integrated Regulation	0.891				
Rewards	-0.213	-0.124			
Pressure	-0.181	-0.166	0.467		
Amotivation	-0.809	-0.903	0.272	0.171	