2016

IDENTIFICATION AND EXAMINATION OF KEY COMPONENTS OF ACTIVE LEARNING

Darrell Scott Kelly
Wright State University

IDENTIFICATION AND EXAMINATION OF KEY COMPONENTS OF ACTIVE LEARNING

A dissertation submitted in partial fulfillment of the
Requirements for the degree of
Doctor of Philosophy

By

DARRELL SCOTT KELLY
B.S. Brigham Young University, 2012
M.S. Wright State University, 2014

2016
Wright State University

WRIGHT STATE UNIVERSITY

GRADUATE SCHOOL

AUGUST 17, 2016

I HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER MY SUPERVISION BY Darrell Scott Kelly ENTITLED Identification and Examination of the Key Components of Active Learning BE ACCEPTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF Doctor of Philosophy.

Dr. Debra Steele-Johnson, Ph.D.
Dissertation Director

Dr. Scott Watamaniuk), Ph.D.
Graduate Program Director

Dr. Debra Steele-Johnson, Ph.D.
Chair, Department of Psychology

Robert E.W. Fyffe, Ph.D.
Vice President for Research and
Dean of the Graduate School

Committee on Final Examination

Dr. Nathan Bowling, Ph.D.

Dr. Dean Parmelee, M.D.

Dr. Valerie L. Shalin, Ph.D.

ABSTRACT

Kelly, Darrell Scott. Ph.D., Industrial Organizational and Human Factors Psychology Ph.D. program, Wright State University, 2016. Identification and Examination of Key Components of Active Learning.

The purpose of this study was to examine key components of active learning. I hypothesized that feedback, accountability, and guided exploration were key components of active learning. I collected survey data from second year medical students (N = 103) in three different active learning interventions: peer instruction (PI), team-based learning (TBL), and problem-based learning (PBL), at six time points. My results did not consistently support my hypotheses. However, I observed a pattern of differences concerning feedback and accountability in the predicted direction in all three interventions. Feedback had a positive effect on professionalism in both PI and PBL, and accountability had positive effects on emotion control and professionalism in both PI and TBL. Also, I found results that raised issues related to each key component. Namely, that perceptions of feedback were influenced by the nature of questions, interactions between individuals, and the source of feedback. Furthermore, accountability was influenced by team membership and a proper measure of guided exploration needs to be developed. This study raised questions regarding which components of active learning affect important outcomes, and what issues affect key components of active learning.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

Active Learning: An Examination of Key Components

Introduction

Active learning is an increasingly popular training strategy in both universities and organizations (Bakker, Demerouti, & Brummelhuis, 2012; Freeman et al., 2014; Saito, Rezende, Falcao, Suzuki, & Gomes, 2014). Many researchers have studied the effects of active learning and have concluded that active learning has positive effects (e.g., Prince, 2004). However, researchers have discussed but not empirically tested key components that might contribute to the positive effects associated with active learning (Burke & Hutchins, 2008). Drawing from prior research on active learning (e.g., Bell & Kozlowski, 2008) and established models of training, I have identified three key components of active learning, namely feedback, accountability, and guided exploration, which I posit contribute to the positive effects of active learning. In the following sections, I discuss active learning and its history, common active learning techniques, the effects of active learning, models of training, and components of active learning interventions.

Active Learning

Bonwell and Eison (1991, p. 2) offered the first widely accepted definition of the term active learning. They stated that active learning is anything that “involves students in doing things and thinking about the things that they are doing.” Active learning is an “umbrella term”. In other words, active learning does not describe a specific technique or intervention. Rather, it is a philosophy that underpins many teaching techniques. Similar to other definitions, definitions and descriptions of active learning have evolved through the years.

History of Active Learning. Many conceptions and designs of learning and teaching are built upon the ideas of constructivism, including active learning (Mayer, 2004). Jean Piaget is regarded generally as the father of the constructivist movement. Though constructivism has been applied in many domains, the underlying premise across domains is that learning is not a passive event. Rather, learning is an active process, and learners need to be active in the process. In constructivist settings, learners are encouraged to seek to construct coherent and organized knowledge (Mayer, 2004).

The ideas of constructivism have a long and rich history. Some researchers have attributed the origins of the constructivism to an 18th century Italian philosopher named Giambattista Vico (Cunningham & Duffy, 1996; Von Glasersfeld, 1989). In the 20th century, thinkers, including Kuhn, Wittgenstein, Dewey, and Rorty, are mentioned as founders of the current definition of constructivism (Cunningham & Duffy, 1996).

Bonwell and Eiason (1991, p. 1) stated that the term active learning appeared in the common vernacular long before the written definition in their book. However, even without using the specific term “active learning”, research involving elements of or ideas related to active learning has a long history. For example, researchers explored the benefits of exploratory and guided learning in the 1950s (Craig, 1953; Craig, 1956; Kittell, 1957). Schermerhorn, Goldschmid, and Shore (1975) researched the effects of student peers teaching each other on learning. Finally, researchers have explored the effects of group dynamics throughout the history of Psychology, including researchers such as William Wundt and Kurt Lewin (e.g., Hogg & Williams, 2000; Lewin, 1947).

The current conception of active learning began its rise in popularity in the 1990s because of research articles that encouraged its use (e.g., Cohn, Atlas, & Ladner, 1994; Dufresne, Gerace, Leonard, Mestre, & Wenk, 1996) and texts that assisted with implementation (e.g., Meyers & Jones, 1993; Silberman, 1996). Iran-Nejad (1990) highlighted the importance of moving away from the fundamental idea that learning occurs through the transfer of information from an external source to a person. He argued that achievement and motivational problems in schools reflected educational systems built on the foundation of learning from externally available knowledge. Further, he argued that learning built on the previously mentioned foundation undermined creativity and limited the scope of what could be learned.

Active learning continues to gain prominence in the eyes of many organizational, educational, and political decision makers. For example, in 2012 the President's Council of Advisors on Science and Technology produced a report detailing how to increase the number of undergraduate college students enrolled in STEM degrees in the United States of America by 33%. In this report, the advisors highly recommended active learning techniques as a way not only to increase enrollment but as a way to improve many issues in undergraduate education, such as student motivation, classroom attendance, and retention of learned material (Olson & Riordan, 2012). The President's Council of Advisors on Science and Technology is an advisory group to the President of the United States of America. The council is consulted frequently and often makes policy recommendations concerning science, technology, and innovation. Also, in this age of technology, active learning techniques are increasingly popular in e-learning interventions (Klasnja-Milicevic, Vesin, Ivanovic, & Budimac, 2011; Koohang & Paliszkiewicz, 2013). To state another example, the medical school in which I conducted this study is completely changing its lectured-based curriculum in order to teach exclusively using active learning techniques. It is doing this due to the positive effects associated with active learning and because of pressure from the Liaison Committee on Medical Education. The Liaison

Committee on Medical Education is the organization that controls the accreditation of medical schools. This committee has strongly advocated for medical schools to increase the usage of active learning techniques. In fact, in some cases this committee has not granted accreditation to schools where they believed that not enough emphasis was placed on active learning techniques. Furthermore, active learning techniques are becoming increasingly popular due to higher education's emphasis on lifelong learning (Sampson & Jackson, 2014).

Characteristics of active learning. According to Bell and Kozlowski (2010), active learning has two fundamental principles that distinguish it from other types of learning. First, active learning approaches provide learners with significant control over their learning. An important feature of any active learning intervention is that learners have a high degree of control over their own learning (Bell & Kozlowski, 2008). The learner has the primary burden for his or her own learning decisions. In non-active learning situations (i.e., passive learning), another individual, not the learner, is in control of the learning. For example, in many classroom settings, students enter a classroom, a teacher lectures for a given amount of time, and then the students leave. Moreover, technology-based learning situations often place the computer is in control of the learning and the learner is a passive recipient.

The second principle that distinguishes active learning from passive learning, according to Bell and Kozlowski (2010), is that active learning adheres to the constructivist vision of learning. The foundational premise of constructivism is that learning is an active process in which learners need to be actively involved in building coherent and organized knowledge (Mayer, 2004). Learners need opportunity to explore the rules, assumptions, procedures, and outcomes of their learning actions (Bell & Kozlowski, 2010). Learners need opportunity for trial and error, to learn from mistakes and success, in order to deeply acquire knowledge. A learner will likely be more invested in his or her own learning when he or she can control how and what he or she learns.

Other researchers have defined active learning differently. For example, Bakker, Demerouti, and Lieke (2012) argued that active learning has three main components. The first component is an individual's *motivation* to start learning something for himself or herself. The second component is that the learner feels in *control* of the learning process. The final component is that the learner can experience a *feeling of mastery* and self-efficacy. Similarly, Prince (2004) defined active learning as any instructional method or training technique that engages participants in the learning process.

Though active learning is not the only philosophy that encourages learners to be actively engaged in their learning, Bell and Kozlowski (2010) suggested that active learning distinguishes itself by incorporating formal training components to help learners shape and understand their learning. Bell and Kozlowski (2008) stated that these formal training components support cognitive, motivational, and emotional processes through which individuals focus their attention, direct their effort, and manage their affect throughout a learning intervention.

Lastly, a central goal of active learning is that active learning interventions seek to influence self-regulatory activity (Kozlowski, Toney, Mullings, Weissbein, Brown, & Bell, 2001). In self-regulatory interventions, students are responsible for their own learning. Self-regulated learners are generally more motivated, have higher levels of self-efficacy, and engage in more meta-cognitive strategies, relative to learners who are not self-regulating (Martens, Gulikers, & Bastiaens, 2004).

Learning and training. The concepts of learning and training are obviously related to and overlap with each other. According to Kraiger and Jung (1997), learning is the extent to which individuals have mastered knowledge, skills, and abilities taught during training. Training has been defined as "the systematic acquisition of skills, rules, concepts, or attitudes that result in improved performance in another environment" (Goldstein & Ford, p.1, 2002). In this project I will refer to training as a process and to learning as an outcome of the training process.

However, learning is not the only outcome that can result from training. Many different outcomes come as a result of different types of training. These outcomes include motivational (e.g., self-efficacy), cognitive (e.g., metacognition), affective (e.g., emotion control), and behavioral factors (e.g., professionalism). I address each of these outcomes in further depth in subsequent sections of this document.

Proposed Research: Key Components of Active Learning

Prior research has informed us that active learning interventions have positive effects on many different variables (e.g., Crouch & Mazur, 2001; Koles et al., 2010; Shin & Kim, 2013), but researchers have not studied empirically the key components of active learning that contribute to the positive effects of active learning. That is, many researchers have focused on validating specific types of active learning interventions, i.e., whether a specific intervention has beneficial effects (e.g., McParland, Noble, & Livingston, 2004; Vasan, DeFouw, & Holland, 2008) but have not focused on the role of different components. Though this prior research is important and has been beneficial, recently researchers have called for the examination of the underlying components of active learning. Freeman et al. (2014) called for what they termed "second generation research." Similar to many researchers, Freeman et al. noted adequate research examining the positive effects of active learning interventions. However, Freeman et al. suggested missing research addressing the reasons why active learning interventions have positive effects. Similarly, Bell and Kozlowski (2010) noted the absence of research providing insight into why active learning has positive outcomes, and called for further research in this area.

Due to the increased interest in active learning interventions and the need to understand the underlying components of active learning that contribute to positive effects, I decided to examine the key components of active learning. In order to determine what the components of active learning interventions could be, I attended many different sessions of three active learning interventions (i.e., Team-Based Learning, Problem-Based Learning, Peer Instruction) over the course of a year. During these sessions I attempted to determine the components of the active learning interventions. From my observations, I created a list of 78 components of active learning. This list revealed two issues. First, this list included too many components to examine in a single study. Second, though I had a list of many components, I did not know which components were most important. Some components could contribute more to the positive effects associated with active learning than other components.

In order to resolve these two issues, I examined the literature related to models of training. I examined training models that had been rigorously studied and widely utilized. These models identified components and processes that are central to achieving positive outcomes in training interventions. While researching these models, I identified key components included in the models that matched with key components that I had identified in the active learning interventions that I observed. Also, these models helped me determine what I believed to be the most important components of active learning. Therefore, on the basis of my literature review, I chose to focus on three key components. These three components were the components that I believed to be the most vital components concerning the explanation of the positive relationships that active learning techniques have with outcomes.

In the following sections, I describe five well-established models that I utilized. These models of training influenced my hypotheses concerning the key components that were important in active learning techniques. Then, I discuss the three key components that I decided to examine in this study which were feedback, accountability, and guided exploration.

Prior Models of Training

Quiñones (1997). Quiñones (1997) modeled contextual factors that influence training effectiveness. Quiñones focused on participation in training design, framing of training, and organizational climate. He argued that these contextual factors are overlooked often in the design and implementation of training. Further, he argued that these contextual factors influence training outcomes (i.e., learning, behavior, results, reactions) and transfer outcomes (i.e., maintenance, generalization) through their effects on trainee characteristics. Also, he discussed trainee characteristics, specifically, motivation to learn, self-efficacy, and fairness perceptions.

Of specific interest to my proposed research project is Quiñones' (1997) theorizing about feedback. Quiñones argued that trainee motivation is influenced by situation factors such as feedback.

Baldwin and Magjuka (1997). Baldwin and Magjuka (1997) modeled pretraining influences on training motivation. They proposed a three-part framework for a pretraining episode: pretraining contextual factors, trainee cognition, and training delivery. Further, the pretraining contextual factors were divided into three sections. The first section was training introduction, which addressed whether the training was voluntary or mandatory, levels of trainee participation, the goals assigned to the training initiative by the trainee, and organizational information (i.e., organizational purpose for training, how objectives contribute to organizational problems, etc.).

Baldwin and Magjuka (1997) argued that these pretraining factors influenced trainees' self-efficacy and outcome expectations, which in turn should influence training outcomes. Specifically, Baldwin and Magjuka (1997) stressed the importance of individual accountability in training environments. Based upon reasoning by Baldwin and Magjuka (1997), I posited that accountability could be another key component to active learning techniques.

Mathieu and Martineau (1997). Mathieu and Martineau (1997) developed a model describing training motivation. They theorized about the individual and situational characteristics that influence pretraining motivation that in turn influence training and work outcomes. Also, they theorized that both individual and situational characteristics moderated the relationship between training outcomes and work outcomes.

The individual characteristics that Mathieu and Martineau (1997) highlighted were demographics, ability levels, education, work experience, manifest needs, personality, goal orientation, job involvement, and career-related attitudes. The situation characteristics that they highlighted were situational constraints, social-psychological influences, and maintenance systems. Feedback is a sub-component of the situation characteristics component of Mathieu and Martineau's model.

Bell and Kozlowski (2010). Bell and Kozlowski 2010 outlined a conceptual model of active learning based on their and others' previous work in the field. Their model consisted of four pieces: training components, self-regulatory processes, individual differences, and training outcomes. Bell and Kozlowski (2010) postulated that training components influenced self-regulatory processes and in turn outcomes. Further, they postulated that individual differences moderated these relationships.

Training components consisted of three categories. The first category was the nature of instruction, which included exploration, experimentation, and inductive learning. The second category was motivation, which consisted of mastery goals, mastery training frame, and goal sequencing. The third category was emotion control, which included emotion-control training, error-management instructions, and attributional retraining.

Further, self-regulatory processes consisted of three categories. The first category was cognition that consisted of metacognition, effortful processing, and mental models. The second category was motivation that consisted of goal orientation, intrinsic motivation, and self-efficacy. The third category was emotion that consisted of anxiety, emotion regulation, and performance attributions.

Thus, Bell and Kozlowski (2010) addressed exploration as an important aspect of instruction, which was one of the training components identified in their model. I posited that the extent to which exploration is guided is an important component in active learning interventions.

Summary. These, models have informed us about important components of training interventions that account for positive training outcomes. However, the key components of these models have not yet been applied to an active learning environment. Based these on prior models of training, I posited that accountability, feedback, and guided exploration are three key components of active learning that relate positive outcomes of active learning.

Key Components

In the following sections, I discuss common components of active learning interventions. The first three sections, accountability, feedback, and guided exploration are components that I examined in my study. I describe each component and discuss how it relates to active learning.

In subsequent sections, I mention briefly other common components and why I chose not to focus on those.

Accountability. Accountability refers to an implicit or explicit expectation that a person might hold relating to the need to justify his or her actions to others (Giessner, Van Knippenberg, Van Ginkel, & Sleebos 2013; Scott & Lyman, 1968). Generally, this expectation will motivate a person to reflect on his or her own decisions and behaviors. There are many different ways to operationally define accountability. In my study I used a measure that examined perceptions of accountability between peers. According to Baldwin and Magjuka (1997), being accountable promotes cooperation, learning, and positive group norms. Accountability has a positive relationship with organizational outcomes. For example, Rohn, Austin, and Lutrey (2003) found that when employees felt accountable for the amount of cash register shortage at the end of the work day, the average daily shortages decreased from $2.27 to $0.06 per day.

Though higher levels of accountability are viewed generally as positive, Hochwarter, Perrewe, Hall, and Ferris (2005) argued for a “sweet spot” to obtain in order to maximize the benefits of accountability. These authors argued that too much accountability might produce negative results because too much accountability will lead employees to perceive lack of autonomy.

Learners in active learning are accountable in many different ways. In many active learning interventions (e.g., Team-based Learning [TBL], Peer Instruction [PI], Problem-based Learning [PBL]), learners are members of a team to whom each member is accountable. Team members are accountable to their teammates for completing the class pre-readings, for contributing to discussion, for demonstrating professionalism as a teammate, etc. Another way learners are accountable is that active learning interventions often require learners to demonstrate knowledge that they have gained (e.g., PI, TBL). Accountability has been shown to have positive effects on outcomes (e.g., Rohn, Austin, & Lutrey, 2003) and is a main component of many active learning interventions.

Feedback. Though feedback has been studied extensively in many contexts throughout the history of Psychology, Ilgen, Fisher, and Taylor (1979) were among the first to offer a conceptual model addressing the effects of feedback on individuals. These researchers defined feedback as "a special case in which a sender or source conveys a message to a recipient" (Ilgen, Fisher, & Taylor, 1979). According to Hattie and Timperley (2007), feedback is information provided by an agent (e.g., teacher, friend, parent, book) regarding aspects of one's performance or understanding. Generally, the goal of feedback is to reduce discrepancies between actual performance and desired performance, i.e., goals (e.g., Hattie & Timperley, 2007). Mathieu and Martineau (1997) argued that feedback has positive effects on trainee motivation and maintenance of what was learned during training.

Feedback can differ in its quality, specificity, detail, quantity, frequency, accuracy, credibility, and orientation (Ilgen et al., 1979). In my study, I examined perceptions of feedback quality between peers. Generally, the more specific the feedback the more utility (Ilgen et al., 1979), especially if the feedback relates to specific goals (Kozlowski et al., 2001). In order to maximize the benefits of feedback, trainers must walk a fine line between giving frequent feedback and giving excessive feedback. Also, feedback should be as accurate and specific as possible.

In active learning interventions, trainees receive various forms of feedback, including peer evaluations, grades, and facilitator feedback. Van Den Bergh, Ros, and Douwe (2013) argued that in many active learning settings teachers or facilitators fail at giving adequate feedback. The many reasons for this failure include: inadequate knowledge of how or when to deliver feedback, improper course design that allows for feedback, and difficulty of giving negative feedback. However, feedback is a main component of many active learning interventions, and research has shown the feedback has positive effects when given properly (e.g., Waldman, Atwater, & Antonioni, 1998).

Guided exploration. According to Debowski, Wood, and Bandura (2001), guided exploration is a predetermined sequence of search activities and responses that have been sequenced in a way as to strengthen the self-efficacy of the trainee. In my study, I examined perceptions of guided exploration. In active learning interventions, guided learning can be sequenced in order to strengthen not only the self-efficacy of the trainee but also the motivation, learning, and transfer of learned information or skills. Guided exploration is not exploration based on the preferences of the trainee or trial and error.

As mentioned above, research has shown that guided exploration has benefits over pure discovery learning. Pure discovery learning enables individuals to learn knowledge without any sort of guidance (Mayer, 2004). Dalgarno, Kennedy, and Bennett (2014) found that trainees in a guided exploration intervention showed significant learning increases over trainees in a pure discovery learning intervention. If learners are given too much freedom during their discovery learning activities, they may fail to come in contact with the material they need to learn (Bell & Kozlowski, 2008). Therefore, discovery learning requires supplemental guidance to help focus the learner’s activities. Bell and Kozlowski (2010) stated that it is ineffective to give learners no guidance during exploratory learning. An appropriate mix of exploration and guidance ensures that learners make proper learning choices and contact the material that they need to learn and understand. Guided exploration is a main component of many active learning interventions, and researchers have shown that guided exploration has positive effects (e.g., Dalgarno et al., 2014).

Other Common Components of Active Learning

The following addresses other components of active learning but components that might be less central to the effectiveness of active learning. Thus, I mention each briefly.

Emotion control strategies. Early research concerning emotion control mainly focused on an individual's ability to manage his or her frustration (Rosenzweig, 1945). However, the definition of emotion control has evolved to include anxiety and other negative emotional reactions (Keith & Frese, 2005). Emotion control is the process through which a person alters and/or controls his or her emotional experience (Melka, Lancaster, Bryant, & Rodriguez, 2011). Generally, these processes are intended to reduce anxiety and negative emotions. These processes greatly vary from person to person.

Participants in active learning interventions need to regulate their emotions. This is because active learning is often difficult and not intuitive. Many people find active learning to be stressful and difficult, especially at the beginning. Being able to control your emotions can help during times of stress and anxiety.

Emotion control strategies are most useful when high attentional and task demands exist (Bell & Kozlowski, 2010). Thus, emotion control is likely to be important in the early stages of skill acquisition and/or when the task is complex or dynamic. For example, emotion control is especially important in the early stages of error management training because that is when errors are most frequent (Keith & Frese, 2005). According to Bell and Kozlowski (2008), components of emotion control interventions include reducing anxiety and frustration and promoting personal control. Also, one of the main goals of emotion control interventions is to reduce state anxiety.

However, one problem with emotion control interventions is that they are often very time consuming. Bell and Kozlowski (2010) stated that a typical intervention could be divided into five or six two hour sessions.

There are many types of emotion control strategies. For example Kanfer and Ackerman (1990) developed a strategy in which they instructed participants to increase the frequency of their positive thoughts and reduce the frequency of their negative emotions. They instructed the participants to not worry or be upset after the commission of errors.

Error framing. Error framing originated from the work of Frese, Brodbeck, Heinbokel, Mooser, Schleiffenbaum, and Thiemann (1991) discussed above. Errors can be framed as positive or negative (e.g., Steele-Johnson & Kalinoski, 2014). Positive error framing encourages individuals to notice errors and to view the errors as helpful and as a tool to learn and better understand a task or situation. Negative error framing encourages individuals to notice errors but to view errors as an inefficiency or as something to avoid (Steele-Johnson & Kalinoski, 2014).

Positive error framing has positive effects. For example, Chillarege, Nordstrom, and Williams (2003) found that positive error framing led to higher test scores. Nordstrom, Wendland, and Williams (1998) found that after participating in error management training, participants had higher levels of intrinsic motivation and decreased levels of frustration.

In many active learning interventions, errors are framed as normal and a valuable component of learning (Gully, Payne, Koles, & Whiteman, 2002). When errors are framed as positive, individuals are encouraged to pay attention and focus on errors, perceiving errors are important and helpful. Individuals are encouraged to learn from errors and use them to improve performance (Steele-Johnson & Kalinoski, 2014).

Time spent on subject. The more time a person spends learning about or thinking about a subject, the greater the learning. Psychologists have tested this notion for over one hundred years (Henderson, 1903; Thorndike, 1908). In their meta-analysis of internet-based learning, Cook, Levinson, and Garside (2010) found that time had a positive relationship with knowledge outcomes.

Time on task is one of the main benefits of active learning. Students spend a significant amount of time on the subject matter outside of class before they engage in additional active learning during class time. In traditional lecture-based classrooms, assigned readings before class lack accountability or motivation for the students to complete the pre-class readings. In most active learning classroom interventions, students are held accountable for their pre-class readings through quizzes or team assignments. Because of this, students in active learning interventions spend more time reading before class. Therefore, active learners spend more time on task than students in lecture-based interventions due to time spent on pre-class readings and actual classroom time.

Decision making.. Research about decision making in psychology dates back to Wundt and Freud (Rothman, 1991). Making better decisions is an obvious outcome that many training programs seek. In medical education, decision appropriateness can be a matter of life or death.

Including a decision-making component into training interventions has positive effects. Sung, Hwang, and Yen (2015) found that students who participated in a training intervention with a decision-making component had greater motivation and problem-solving skills than a control group. However, other components contaminated the training intervention, so one cannot conclude that the sole reason for the positive effects was the decision-making component. Active learning interventions frequently include components that require trainees to make decisions including individual tests, group tests, and group decisions resulting from group discussion.

Repetition. The effects of repetition on learning is another topic that has a long history in Psychology (e.g., Smith, 1896). The general conclusion from the many experiments in this area is that repetition has positive effects on learning in most situations (e.g., Dewhurst & Anderson, 1999; Salasoo, Shiffrin & Feustel, 1985). Repeated exposure to material is one advantage that active learning techniques have over other traditional teaching techniques. Students in active learning interventions are exposed to the same material multiple times in multiple contexts. For example, in a TBL setting, students study certain material outside of class. Then, they are tested individually on that material, tested as a team on the same material, and finally spend class time applying the material to real world situations.

Memorable Experiences. Active learning interventions create more memorable learning experiences than traditional teaching methods. For example, in PI interventions, students are presented with a case containing usual circumstances. Then, the case unfolds over multiple questions. Not only is the case memorable, but because the case unfolds over time it becomes an even more memorable experience.

Graded vs. ungraded. Research and opinion concerning the issue of whether to grade student responses during certain active learning activities remains mixed. Certainly, many schools and programs adhere to the need to assign grades for class activities and tests, but many programs have adopted a pass/fail approach and espouse that their goal is for students to learn the material and not to learn for a grade.

Freeman et al. (2007) found that students performed better on in-class questions that were graded vs. ungraded. In contrast, Shin, Haynes, and Johnston (1993), found that students in an ungraded curriculum at one university outperformed students in a graded curriculum at another university. Furthermore, after a three-year study, Hopkins, Oldridge, and Williamson (1965) found no significant differences between students in graded vs. ungraded classrooms. Further adding to the debate between graded vs. ungraded classrooms is our knowledge of mastery and learning goal orientation. Therefore, more research is needed to determine the effects of grades on training outcomes.

Outcomes of Active Learning

Researchers have studied many different active leaning strategies. There is a large body of evidence that leads to the conclusion that active learning interventions generally have positive outcomes. Often, research into active learning techniques has contrasted active learning with traditional lectures, in which students are passive recipients of instruction from a teacher (Prince, 2004). In the following sections, I elaborate on the outcomes of active learning. The first five sections, learning, metacognition, self-efficacy, emotion control, and professionalism represent outcomes that I am interested in examining in this study. I chose these outcomes because both researcher and practioners have established these outcomes as important in active learning training environments (e.g., Huffaker & Calvert, 2003; Kozlowski & Bell, 2010). I sequence my discussion of these five outcomes by first defining the outcome and discussing previous research findings concerning the outcome and active learning. Next, I discuss how each of the proposed key components of active learning (i.e., accountability, feedback, guided exploration) are related to each outcome based upon previously stated theoretical reasoning. Then, I discuss research results related to the proposed key components and outcomes. After each section, I state my hypotheses. The remaining sections address outcomes that other researchers have examined in connection with active learning.

Primary Outcome Variables

In the following sections, I describe the five outcomes that I hypothesized would be influenced by active learning interventions. These sections are ordered according to Kraiger, Ford, and Salas' (1993) model of training evaluation, which organizes training outcomes into three categories; cognitive-based, affective-based, and skill-based. The outcomes that I categorized as cognitive-based are learning and metacognition. The outcomes that I categorized as affective-based are self-efficacy and emotion control. The outcome that I categorized as skill-based is professionalism.

Learning. As the name implies, active learning techniques are designed to influence learning. In addition to the examples cited above, a large body of literature has supported the hypothesis that active learning has positive effects on learning. For example, Yoder and Hochevar (2005) found that students performed better on test questions when students were taught using active learning techniques. Similarly, McCarthy and Anderson (2000) found that students in active learning classrooms did better on standard evaluations than students in traditional lecture classrooms. Freeman et al. (2014) found that though active learning was effective in all different class sizes, active learning techniques had the greatest impact on student performance in class sizes of less than fifty. Keeler and Steinhorst (1995) found that students engaged in active learning statistics courses had higher class scores and were less likely to drop the course compared to students in traditional lecture statistics courses. McConnell (1996) found that students in active learning classrooms outperformed students who were not in active learning classrooms in the domain of computer science. Also, Freeman et al. (2007) found that an active learning classroom in the domain of Biology had lower failure rates, higher total exam points, and higher midterm exam scores on an identical. Therefore, research shows that active learning interventions have positive relationships with learning.

Furthermore, the components (i.e., accountability, feedback, guided exploration) that I propose are key to active learning interventions might have positive relationships with learning. Newmann, King, and Rigdon (1997) argued that feeling accountable to someone has a motivating effect that results in individuals exerting more effort to achieve goals. Therefore, as students feel more accountable, they might apply more effort to their schoolwork and learning could increase. Concerning feedback, Hounsell (2007) stated that feedback has a long history of positively influencing learning in both formal and non-formal educational settings. He stated that this effect occurs because when a person receives feedback he or she has a clear sense of how well he or she is doing and what he or she needs to do to improve. This allows the learner to overcome weaknesses and become more efficient during the learning process. Also, guided exploration might have a positive relationship with learning. This might occur because learning is an active process in which learners attempt to construct and organize knowledge (Mayer, 2004). As mentioned previously, guided exploration allows learners to construct and organize their own knowledge that should lead to increases in learning.

In addition to theoretical rationale, the hypothesis that the key components of active learning should have positive relationships with learning is strengthened by empirical research. In his review of cooperative learning, Slavin (2011) showed that in 37 of 44 studies individual accountability levels had a positive relationship with learning. Also, in their meta-analysis, Azeyedo and Bernard (1995) found that feedback had a positive relationship with learning. Finally, Debowski, Wood, and Bandura (2001) showed that guided exploration has positive relationship with performance and learning. However, none of this research has examined these three components as part of an active learning intervention. Given these empirical results and previously stated theorizing, I hypothesized the following.

Hypothesis 1: Feedback, accountability, and guided exploration within an active learning intervention will have positive relationships with learning.

Metacognition. The term metacognition emerged in the 1970s from work by Brown (1975) and Flavell (1979). Frequently, metacognition is defined as the degree of knowledge about and control of one's own cognitions (Brown, 1975; Flavell, 1979; Palincsar & Brown, 1987). This includes planning, monitoring, and changing to meet appropriate goals (Ford et al., 1998). Metacognition is different from cognitive ability, and research has suggested that metacognitive skills are an even more important predictor of training performance than cognitive ability (Kozlowski et al., 2001). Metacognition is important in training environments because individuals with high levels of metacognition perform better during training. Also, during training a person could be thinking about the task, how to improve his or her task performance, how to use these skills to impact future goals, and how he or she is performing during the training (Kozlowski et al., 2001).

One reason why active learning is becoming increasingly popular is because universities and organizations realize that students need to develop higher order metacognitive skills in order to keep up with our ever-changing society (Van Den Bergh, Ros, & Beijaard, 2013). People are bombarded with more information today than they ever have been in the history of the world, and higher order metacognitive skills are increasingly vital. Metacognitive skills are developed in active learning environments because as students become more responsible for their own learning they develop the self-regulatory processes associated with metacognition (e.g., knowing what you know, knowing how you learn, self-evaluation). Many activities that are part of active learning interventions help students develop metacognitive skills, such as guided exploration, group discussion, and goal setting. Linton, Pangle, Wyatt, Powell, and Sherwood (2014) argued that writing and discussion promote the development of metacognition. Bell and Kozlowski (2008) argued that individuals engage in various levels of metacognition during active learning interventions.

Metacognition is an important outcome in many training interventions (Bell & Kozlowski, 2010), and generally active learning interventions increase levels of metacognition. Keith and Frese (2005) found that an active learning intervention (i.e., error management training) stimulated more metacognition than a non-active learning intervention. Bell and Kozlowski (2008) found similar results. These authors found that guided exploration increased levels of metacognition over proceduralized instruction intervention. Even different active learning interventions can produce different levels of metacognition in learners. For example, Stefanou, Stolk, Prince, Chen, and Lord (2013) found that students in two courses experienced increases in metacognition. Further, students who were enrolled in a project-based course reported higher levels of metacognition than students in a problem-based course. Research has shown active learning interventions have positive effects on metacognition.

Though there is little research concerning the direct relationships of the proposed key components with metacognition, there is ample theoretical support. For example, Lerner and Tetlock (1999) stated that a person who feels accountable feels the implicit or explicit need to justify his or her beliefs, feelings or actions. This relates to metacognition because in order for a person to be able to justify his or her beliefs he or she must be aware of what he or she believes. Therefore, as a person becomes more aware of the need to justify himself or herself, he or she might attempt to become more aware of what he or she is thinking. Flavell (1979) stated that an important part of metacognitive knowledge is learning which tasks are demanding or difficult. Feedback likely plays a key role in helping an individual understand their level of competency on a task, which might lead to an understanding of how demanding or difficult a task is. Because of this, Paris and Winograd (1990) suggested that instructors should offer feedback in order to promote higher levels of metacognition among their students. Finally, guided exploration might have a positive influence on metacognition. Sweller, Mawer, and Ward (1983) found that an important antecedent to the development of metacognition is the opportunity to engage in self-directed learning. Guided exploration is a form of self-directed learning that should result in the development higher levels of metacognition (Bell & Kozlowski, 2008).

As mentioned previously, little research has focused on the effect of accountability on metacognition, although Jahromi and Mosallanejad (2014) found that accountability had a positive relationship with metacognition. More research has been done concerning the relationship between feedback and metacognition. For example, Brady, Seli, and Rosenthal (2013) found that feedback had a positive effect on metacognition. Less research has addressed the relationships between guided exploration and metacognition. Yet, Bell and Kozlowski (2008) found that exploratory learning, which as discussed above is similar to guided exploration, had a positive relationship with metacognitive activity. However, none of this research has examined these three components as part of an active learning intervention. Therefore, given the previous research results and the theory discussed above, I hypothesized the following.

Hypothesis 2: Feedback, accountability, and guided exploration within an active learning environment will have positive relationships with metacognition.

Self-efficacy. Self-efficacy is a person's belief about his or her ability to control his or her level of functioning and events that are affecting his or her life (Bandura, 1977, 1997). A person's self-efficacy levels affect how much effort and energy he or she put into activities, his or her goals, and his or her motivation to complete tasks. Over the years, researchers have examined self-efficacy as a dispositional trait and as something that is task specific (Gist & Mitchell, 1992). Sampson and Jackson (2014) argued that active learning interventions have a positive effect on self-efficacy.

Self-efficacy can be increased through training. Providing easy, proximal, and specific goals and influencing the trainees in a way that makes them feel that they have control over their training benefits self-efficacy (Kozlowski et al., 2001).

Not all active learning interventions include the components discussed above that can increase self-efficacy. However, there is some evidence that active learning interventions have a positive influence on self-efficacy. Bell and Kozlowski (2008) found an active learning intervention influenced metacognition and in turn self-efficacy. Also, they found that state prove goal orientation mediated the positive relationship between error framing and self-efficacy and that state anxiety mediated the positive relationship between emotion control and self-efficacy. Similarly, Griffin and Griffin (1998) found that an active learning intervention had a positive effect on self-efficacy. However, more research is needed to establish firmly the positive effects of active learning interventions on self-efficacy.

Theoretical bases support the idea that the key components of active learning should have positive relationships with self-efficacy. Accountability should have a positive relationship with self-efficacy because individuals who are accountable are more likely to conduct a self-evaluation of their performance (Royle, Hall, Hochwater, Perrewe, & Ferris, 2005). This is because accountable individuals feel the need to justify their actions to those around them (Lerner & Tetlock, 1999) and therefore likely preemptively conduct a self-assessment. Individuals who conduct a self-assessment are more likely to increase their self-efficacy because individuals typically have a self-serving positivity bias (Mezulis, Abramson, Hyde, & Hankin, 2004). This positivity bias may cause an individual to self-assess himself or herself positively which should lead to increased self-efficacy. Providing feedback, particularly about successes, might lead individuals to believe that they succeeded due to their ability levels (Schunk, 1983). This likely positively affects self-efficacy. Furthermore, guided exploration should have a positive relationship with self-efficacy for similar reasons. Individuals in a guided exploration intervention succeed due to their own ability to categorize and learn the proper information (Debowski, Wood, & Bandura, 2001). They recognize the reason for their successes, which should lead to increased self-efficacy.

However, more empirical research is needed to determine outcomes of the components of active learning on self-efficacy. Few experiments have been conducted examining accountability and self-efficacy, and these experiments have provided inconclusive results. For example, Royal, Hall, Hochwarter, Perrewe, and Ferris (2005) found a non-significant relationship between accountability and job self-efficacy. Concerning feedback and self-efficacy, Schunk (1982) found that feedback had a positive relationship with self-efficacy. Similarly, Debowski, Wood, and Bandura (2001) found that guided exploration had a positive effect on self-efficacy. However, these relationships have not been examined in an active learning environment. Therefore, given the previous research results and the theory discussed above, I hypothesized the following.

Hypothesis 3: Feedback, accountability, and guided exploration within an active learning environment will have positive relationships with self-efficacy.

Emotion control. Keith and Frese (2005) defined emotion control as a skill that an individual uses to keep performance anxiety and other negative emotions under control while that person is engaged in a task. Emotion control strategies are individual techniques that a person uses to manage his or her emotional states (e.g., suppression, reevaluation). Bell and Kozlowski (2008) found that emotion control strategies led to a decrease in state anxiety. Keith and Frese (2005) found that emotion control mediated the relationship between error management training and transfer performance. Sitzmann and Ely (2011) argued that emotion control can help improve learning because a person who is high in emotion control can focus on the task and block out distractions better than a person low in emotion control. However, Stizmann and Ely (2011) found that emotion control had no significant effects on learning. In sum, more research is needed concerning the predictors of emotion control.

Accountability should have a positive relationship with emotion control. This should occur because an individual who is high in accountability feels social pressure to justify his or her actions. Emotional responses, especially negative emotional responses, are not socially acceptable (Baumeister, 1982). Therefore, an individual who is high in accountability should be influenced by social pressure to control his or her emotions, which should lead to an increase in emotion control. Also, feedback should have a positive relationship with emotion control in an active learning environment. The reason for this is simple. Individuals in active learning interventions often receive feedback regarding their social interactions with their teammates, peers, etc. This type of feedback should positively influence emotion control. Finally, guided exploration should have a positive relationship with emotion control. Keith and Frese (2005) argued that error management training, a type of active learning intervention of which guided exploration is a large component, should have beneficial effects on emotion control due to the nature of the intervention. In an intervention with guided exploration, trainees inevitably go through the process of making mistakes and learning from those mistakes. As this process continues, the trainee learns to control his or her emotions when mistakes occur in order to overcome them. This should have a positive effect on emotion control.

Little research has examined the relationships of the key components of active learning with emotion control. To my knowledge, no published research has examined the relationship between accountability on emotion control. Similarly, to my knowledge no published research has examined the relationship between feedback on emotion control. However, Keith and Frese (2005) found that a training intervention high in guided exploration had positive effects on the trainee's level of emotion control. Yet, they did not directly examine the effects of guided exploration on emotion control. Nevertheless, given the previous research results and the theory discussed above, I hypothesized the following.

Hypothesis 4: Feedback, accountability, and guided exploration within an active learning environment will have positive relationships with emotion control.

Professionalism. Professionalism is an important element in the topic of this research—medicine. The American Board of Internal Medicine defined professionalism as commitment to: the highest standards of excellence in the practice of medicine and in the generation and dissemination of knowledge, sustain the interest and welfare of patients, and be responsive to the health needs of society (Chisholm, Cobb, Duke, McDuffie, & Kennedy, 2006). Few researchers have studied the relationships between active learning on professionalism. Those few studies that have examined the influence of active learning on professionalism do not support a hypothesis that active learning interventions have an effect on professionalism. For example, Beatty, Kelley, Metzger, Bellebaum, and McAuley (2009) found a non-significant effect of TBL on professionalism. Also, Wimmers and Lee (2015) found that PBL did not have a significant effect on professionalism. Based on these results, it is reasonable to reach a preliminary conclusion that active learning does not have a direct relationship with professionalism.

Also, the literature does not support the conclusion that the key components of active learning should be related to professionalism. There are no research articles examining the relationship between accountability and professionalism. There is research examining the relationship between feedback and professionalism, but the literature has provided mixed results. For example, Brinkman et al. (2007) found that based on other source ratings, feedback had a significant positive effect for ratings completed by one group and a non-significant effect for ratings completed by another group. As with accountability, there are no experiments that test the relationship between guided exploration and professionalism. Given the limited and mixed research results and theory discussed above, I posed the following research question.

Research Question 1: Do feedback, accountability, and guided exploration within an active learning environment have relationships with professionalism?

This study was conducted over the course of an academic year. Thus, the data I collected was longitudinal in nature. Due to this, I was interested in examining the relationships between time and the key components variables and outcome variables. Though I did not have any a priori hypotheses concerning these relationships, I conducted analyses that tested these relationships with time. I collected qualitative data in the form of participant interviews to aid in the interpretation of my results. I discuss these analyses and results further in the method and results sections of this document.

Other Variables Influenced by Active Learning

In the following sections, I describe nine other outcomes that might be influenced by active learning interventions. Though I chose not to examine these outcomes in my study, I have included these outcomes in this section because learning about these outcomes helped me understand more about active learning, and helped me decide what to examine in my study. These sections are ordered according to Kraiger, Ford, and Salas' (1993) model of training evaluation, which organizes training outcomes into three categories; cognitive-based, skill-based, and affective-based. The outcome that I categorized as cognitive-based is lifelong learning. The outcomes that I categorized as skill-based are complex skill development, performance, and adaptive transfer. The outcomes that I categorized as affective-based are intrinsic motivation, goal orientation, perceived utility, anxiety, and peer support.

Cognitive outcome: Lifelong learning. The term lifelong learning evolved from the adult learning literature (Lily, 1951) during the 1960s. Lifelong learners are self-motivated, are able to identify their own learning needs, and evaluate strategies that will help them increase their learning (Husen, 1968; Stefanou, Stolk, Prince; Wolfe, 1963). This definition highlights the similarities between lifelong learning and the goals of self-regulated learning. Researchers have argued for the benefits of active learning on lifelong learning (Miflin, Campbell, Price, & Mirlin, 2000; Phil, 2000). However, no studies have examined lifelong learning as an outcome of active learning interventions.

Skill-based outcome: Complex skill development. Active learning techniques have a positive influence on the acquisition of complex skills. For example, Martin, Rivale, and Diller (2007) found that students in active learning interventions showed greater improvement in innovative thinking abilities as compared to students in traditional lecture based classes. Also, Thompson, Califf, and Mooney (1999) found that an active learning intervention significantly reduced the amount of instruction that individuals needed to achieve a given level of performance on a complex language task.

Usually active learning interventions are used to teach complex skills. This is because active learning is unlikely to have any benefit over passive or proceduralized approaches to training when the goals of the training are basic declarative knowledge and skills (Bell & Kozlowski, 2010). One reason for this could be that the acquisition of basic skills does not require self-regulated learning by the trainee; it could require other factors such as repetition, memorization, etc.

Skill-based outcomes: Performance. Freeman et al. (2014) conducted a meta-analysis of the effects on student performance of active learning interventions vs. traditional learning interventions (i.e., lecture) in science, technology, engineering, and math courses. They found a weighted mean difference of .47 in favor of the active learning interventions. Also, these researchers found that failure rates increased by 55% in traditional learning interventions vs. active learning interventions (lecture failure rates = 33.8% and active learning failure rates = 21.8%). Additionally, Freeman et al. (2014) found that discipline moderated the effects of active learning interventions. They concluded that active learning is effective across science, technology, engineering, and math disciplines.

Researchers have observed similar benefits of active learning outside of classroom settings. Bakker, Demerouti, and Brummelhuis (2012) found that active learning was positively related to task performance and contextual performance. Also, Taris, Kompier, De Lange, Schaufeli, and Schreur (2003) found that active learning had positive effects on employees in high control/low demand jobs. Furthermore, Naveh Katz-Navon, and Stern (2015) found a negative relationship between active learning environment and employee errors.

Skill-based outcomes: Adaptive transfer. Research concerning adaptive transfer was conducted as early as the 1970s (e.g., Mayer & Greeno, 1972). Today, adaptive transfer is defined as gaining knowledge in one situation and applying it to generate a solution to a problem in a novel situation (Hatano & Inagaki, 1986; Kluge, Sauer, Burkolter, & Ritzmann, 2010). Active learning techniques should contribute to adaptive transfer because trainees in active learning interventions often are asked to generate solutions to novel problems or overcome unexpected problems. There is evidence that active learning interventions lead to higher rates of transfer, including adaptive transfer. For example, Bell and Kozlowski (2008) found that self-evaluation activity positively influenced strategic knowledge which in turn influenced adaptive transfer. Kozlowski, Gully, Brown, Salas, Smith, and Nason (2001) found that trainees in a mastery training intervention had increased levels of adaptive transfer.

Analogical transfer is the transfer of knowledge and/or skills to situations that are similar in structure but not content relative to training (e.g., Kim & Lee, 2001). Research has shown that active learning does not always lead to positive analogical transfer but frequently leads to positive adaptive transfer. For example, McDaniel and Schlager (1990) concluded that discovery learning did not provide any benefit to analogical transfer but benefited adaptive transfer. Similarly, researchers have found that exploratory learning (Bell & Kozlowski, 2008) and mastery training (Kozlowski et al., 2001) led to positive adaptive transfer.

Affective-based outcome: Intrinsic motivation. Researchers have defined intrinsic motivation as a person’s inherent tendency to seek new things and challenges, to expand one’s own capacities, and to explore and learn (Deci, 1971; Ryan & Deci, 2000; White, 1959).

Benware and Deci (1984) conducted an experiment in which one group experienced an active learning intervention and the other group experienced a passive learning intervention. The active learning condition induced higher levels of intrinsic motivation than the passive learning condition. Similarly, Young (2005) found that an active learning environment had indirect effects on intrinsic motivation through goal orientation, perceived autonomy, and perceived competence. Furthermore, Rawsthorne and Elliot (1999) found that mastery goal orientation was related to high levels of intrinsic motivation.

Affective-based outcome: Goal orientation. Nicholls (1975) and Dweck (1975, 1986) independently conceptualized goal orientation in the 1970s. An individual's goal orientation is a person's dispositional or situational goal preferences in achievement situations (Dweck, 1975; Payne, Youngcourt, & Beaubien, 2007). Goal orientation is characterized as a somewhat stable individual factor. However, situational influences can affect goal orientations (Button, Mathieu, & Zajac, 1996; Heinz & Steele-Johnson, 2004). Usually researchers have examined two goal orientation dimensions, i.e., performance and mastery goal orientation, and sometimes further separated performance orientation into approach and avoid dimensions.

Performance goal orientation refers to the degree to which individuals are focused on performing for people who are observing them (Vandewalle, 1997). Individuals high in performance orientation want to exceed set standards and outperform others (Ford et al., 1998). These individuals often avoid challenging tasks because they are afraid to fail in front of others. Individuals high in performance orientation believe that their abilities are not malleable (Kozlowski et al., 2001).

Mastery goal orientation refers to the degree to which an individual desires to develop competence and master tasks (Vandewalle, 1997). Further, Kozlowski and Bell (2006) stated that mastery orientation is positively associated with intrinsic motivation, challenge seeking, persistence, self-efficacy, and metacognitive ability. Individuals high in mastery orientation desire to learn new skills, thoroughly understand tasks, and set goals that allow them to achieve mastery in different areas (Ford et al., 1998). Research has shown that mastery orientation has positive effects in workplace and classroom settings. Students with high levels of mastery orientation develop more effective learning strategies, have more positive attitudes towards class, and believe that they can succeed based on their effort (Ford et al., 1998). Individuals high in mastery orientation maintain higher motivation throughout training (Kozlowkski et al., 2001) and have higher rates of learning and transfer (Bell & Kozlowski, 2008).

Affective-based outcomes: Perceived utility. Utility was first measured by Greller (1978; Dusterhoff, Cunningham & MacGregor, 2014). Training is perceived to be high in utility if an individual believes the training will lead to a desired outcome (Chiaburu & Lindsay, 2008). Perceived utility was a variable that Burke and Hutchins (2007) identified as an important component of training transfer. Dusterhoff, Cunnigham, and MacGregor (2014) found that utility had a positive relationship with performance appraisal satisfaction.

Affective-based outcome: Anxiety. The term anxiety has existed for hundreds of years, and research pertaining to anxiety dates back to some of the fathers of modern Psychology, namely Freud and Pavlov (Klein, 2002). As it is defined today, anxiety is an unpleasant state of distress and/or psychological arousal due to some stimulus (Kouchaki & Desai, 2014). Many researchers have associated anxiety with stress, nervousness, and/or dread (e.g., Gray, 1991). All humans experience anxiety, and generally when a person experiences anxiety, the person is highly motivated to reduce the anxiety. Anxiety has both dispositional and state components. Trait anxiety is a person's natural level of anxiety, and state anxiety is a more temporary condition. Generally, anxiety has a negative effect in most situations, especially training settings, because once a person experiences anxiety, the anxiety diverts a person's attention and motivation away from the task.

Because active learning often is used to train complex knowledge and skills, active learning interventions frequently induce anxiety. Active learning researchers often have incorporated emotion control strategies in the training in an attempt to reduce anxiety (Bell & Kozlowski, 2008). However, research has not examined the direct effects of active learning interventions on anxiety.

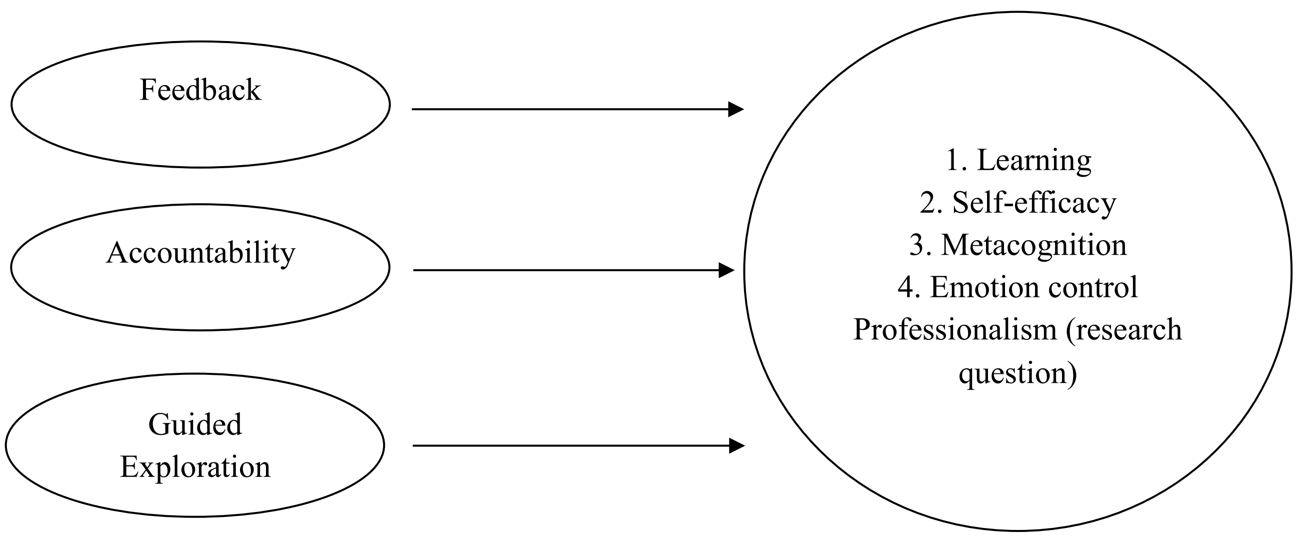
Affective-based outcomes: Peer support. Peer support emerged from the social support literature in the 1980s (Westman, Eden, & Shirom, 1985). Peer support is social support from immediate coworkers (Westman, Eden, & Shirom, 1985). It is characterized by respect, shared responsibility, and mutual agreement of what is helpful (Mead, Hilton, & Curtis, 2001). In their review of the training transfer literature, Burke and Hutchins (2007) identified peer support as an important feature that promotes training transfer. Peer support encourages a trainee to apply newly trained skill in the work context (Massenberg, Spurk, Kauffeld, 2015).

Delayed or negative active learning effects. One of the most common findings is that the results of active learning are often not immediate. In fact, individuals often perform worse at the end of an active learning intervention as compared to a passive learning intervention. For example, Bell and Kozlwoski (2008) found that individuals who engaged in exploratory learning performed worse at the end of the training session than those who were involved in proceduralized learning. Keith and Frese (2005) found that trainees in error training interventions had lower levels of performance due to the high frequency of errors. Also, trainees often arrived at the wrong conclusions in error training interventions and exploratory learning interventions. However, the goal of many active learning interventions is to increase performance after the training intervention, and as has been described above, that goal is usually met.

Conclusion. Active learning strategies are diverse and include many different components. I have chosen accountability, feedback, and guided exploration as three components that I posit relate to the positive effects of active learning. I have based my reasoning on prior theoretical work and empirical work research. A summary of my hypotheses is in Figure 1. My predictions focus on relationships between primary predictors and primary outcomes.

Figure 1

Predictors and Outcomes of Active Learning



Common Active Learning Techniques

In the next section, I describe some of the most common techniques that fall under the active learning umbrella. The first three techniques that I describe, namely problem-based learning, team-based learning, and peer instruction, are the three active learning techniques that I examine in this study. I describe other popular active learning techniques as well but in less detail.

Problem-based learning. PBL is a method of active learning that is very popular in medical education and becoming increasingly popular in other classroom settings as a replacement of lecture-based instruction (Hung, Jonassen & Liu, 2008; Savery, 2015). McMaster University is credited generally with the creation of PBL in the 1960s (Neufeld & Barrows, 1974). When it was developed, there was no philosophical or theoretical reasoning underpinning its structure (Neville, 2009). However, the influence of PBL began to accelerate in the United States in the 1980s largely due to the Report of the Panel on the General Professional Education of the Physician and College Preparation for Medicine, which highly recommended the use of PBL in medical education (Hung, Jonassen, & Liu, 2008). The growth of PBL continues to this day.

According to Dolmans, Michaelsen, Van Merrienboer, and Van Der Vleuten (2015), PBL has four main components: learning occurs through solving problems (sometimes called cases), students meet in small group sessions, there is a facilitator present in the small group sessions who can offer varying levels of input, and learning also occurs outside the group setting in the form of self-study. PBL is designed to improve critical thinking and problem solving skills and to help students apply theory to practice (Shin & Kim, 2013). In a medical school setting, students are assigned randomly to small groups and are given a "case." Generally, the cases revolve around a patient for whom the students need to diagnose and identify treatment. The idea is that the case gives the details of the patient, and as a group the students identify relevant topics (e.g., the impact of a certain drug on pregnancy, the causes of arthritis) that they need to study in order to know what is happening with the patient. Then, students study the topics individually and later meet together with their small group to discuss their findings.

In their meta-analysis of the PBL literature, Shin and Kim (2013) found a positive effect (Cohen's *d* = 0.70) for PBL's effect on student learning. Also, they found that PBL had positive effects on satisfaction with training and skill development. Kong, Qin, Zhou, Mou, and Gao (2014) found evidence that PBL helped nursing students improve their critical thinking skills. Other researchers have lamented that there is a lack of evidence supporting the benefits of PBL (e.g., Sanson-Fisher & Lynagh, 2005; Savery, 2015). These researchers argued that many of the studies that have examined PBL are methodologically flawed. They noted the lack of control groups and appropriateness of using meta-analyses for evaluating medical school curricula (Sanson-Fisher & Lynagh, 2005). Because of this, they argue that the conclusions of much of the research concerning PBL are not valid. However, despite this lamentation, PBL's popularity has continued to increase.

Team-based learning. TBL is another active learning method that is increasingly popular in university settings. TBL is a structured technique that involves three steps. These steps are: pre-class readings, readiness assurance process, and application learning activities (Michaelsen, Sweet, & Parmelee, 2008). Students are required to complete pre-class readings to prepare for class discussion. The readiness assurance process consists of four parts: an individual test, a team test that is exactly the same as the individual test, a team appeals process, and the instructor briefly answering questions about the test. After the tests, most of the class time is devoted to the application learning activities. In addition to describing the structure of TBL, Michealsen, Sweet, and Parmelee (2008) identified four essential elements of TBL, which are: properly formed and managed groups, students being held accountable for their individual and group work, frequent and timely feedback, and group assignments that promote learning and team development.

Though learning in teams is not a new concept, with research in the 1960s addressing this in terms of stages of group development (e.g., Tuckman, 1965), the technique described above was developed in the 1970s to overcome the problem of increased enrollment in colleges of business administration (Michaelsen, Watson, Cragin, & Fink, 1982). TBL enabled large classes to meet together but still have learning occur in a small group setting. This decreased cost for universities and also improved student learning outcomes (Michaelsen, Watson, Cragin, & Fink, 1982). TBL was used first in a medical school setting in 2001 at Baylor University (Haidet, O'Malley, & Richards, 2002). Since its introduction, TBL has spread to many universities and has been used in a variety of disciplines throughout the world (Sisk, 2011).

Koles, Stolfi, Borges, Nelson, and Parmelee (2010) found that students performed better on exam questions if the material in the exam question was taught using TBL versus other teaching methods. Also, these authors found that students in the lowest quartile of the class gained more benefit than those in the highest quartile. Moreover, TBL impacted student attitudes. Vasan, DeFouw, and Compton (2009) found that students had a favorable perception of TBL that was independent of the grades that they earned in the TBL course. Furthermore, Chung, Rhee, Baik, and Oh-Sun (2009) found that students perceived TBL to be more engaging, effective, and enjoyable than conventional teaching techniques.

Peer instruction. PI was developed by Eric Mazur in 1991 because he was unsatisfied with the learning that was taking place in his lecture-based physics course. Though PI is not as rigidly structured as TBL, there are basic elements present in a PI course. Students are required to complete pre-class readings in order to prepare for class discussions. The classroom sessions are divided into short presentations in which the instructor usually lectures briefly and then asks the students a question. Students are given a short time to answer and report their answers. Usually, if a certain percentage of students do not answer correctly, time is devoted to student discussion. In this discussion, students discuss their answers with those seated around them and try to convince their peers of the correctness of their answers. At the conclusion of the discussion, the students report their answers again, and the professor explains the answer and moves on to the next topic (Crouch & Mazur, 2001).

PI has positive outcomes for students. Based on ten years of data, Crouch and Mazur (2001) found that student test scores improved after implementation of PI. Crouch, Watkins, Fagen, and Mazur (2007) reported student gains in both conceptual reasoning and quantitative problem solving. Zingaro (2014) found that PI positively influenced self-efficacy.

Error management training. Although not a focus in my dissertation research, error management training is another popular active learning technique. The term “error management” was developed by Frese, Brodbeck, Heinbokel, Mooser, Schleiffenbaum, and Thiemann (1991). These authors did not agree with the commonly held notion that errors are bad in training environments. They stated four reasons why errors could play a positive role in training. First, they argued that a person’s mental model of a system is enhanced when a person makes an error. Second, they posited that a person’s mental models are of a higher quality when the mental models include potential errors or problem areas. Third, if a trainer desired to have an error-free environment, the trainer might restrict the types of strategies or learning techniques of the trainees. Fourth, Frese et al. (1991) correctly observed that errors naturally appear in the work environment. Generally, it is difficult to eliminate all errors from a work situation.

Usually errors are viewed as negative, and because of this individuals try to avoid errors whenever possible. However, the fundamental goal of error management training is to encourage errors in order to learn from them. Error management interventions have three distinguishing features according to Keith and Frese (2008). One is that participants are given very minimal guidance, and participants are encouraged to test and explore to accomplish goals. Second, trainers create an environment in which errors are highly likely to occur. Third, participants are explicitly instructed or encouraged to make errors. Participants are informed about the positive effects of errors, and every effort is made to positively frame errors during the intervention.

In error management training, tasks are usually difficult, which increases the likelihood that errors will be committed. As an example, in a study conducted by Keith and Frese (2005), participants were given a small amount of time and instructed to exactly replicate a complex PowerPoint slide set with minimal PowerPoint training beforehand.

Research has shown that error management training has positive effects. For example, Gully, Payne, Koles, Whiteman (2002) found that error management training has a positive effect on self-efficacy, knowledge, and performance. Furthermore, in their meta-analysis of the error management literature, Keith and Frese (2008) found that the mean effect of error management training was positive (Cohen's $d = 0.44$).

Exploratory learning and guided exploration. Exploratory learning, also called discovery learning, is a technique that asks individuals to seek information or solve a problem, usually individually, with very few or in many cases no outside influences guiding the exploration. For example, in a classroom setting, students could be given a complex problem with the instructions to find the solution before the next class. These were the only instructions that the students would receive. This technique was promoted as early as the 1950s and 1960s (Bruner, 1961; Tuovinen & Sweller, 1999). Guided exploration is the same as exploratory learning except instead of receiving no guidance, the trainees receive direction, hints, feedback, etc., in order to help them complete the goals of the training (Mayer, 2004).

Many researchers have argued that pure exploratory learning without any guidance is less beneficial than guided exploration (e.g., Kirschner, Sweller, & Clark, 2006). For example, in his review of exploratory learning, Mayer (2004) stated that there is enough research to conclude that exploratory learning has very little benefit over guided learning. One of the main reasons for this conclusion is that in an exploratory learning intervention there is no guarantee that the individual will encounter the relevant rules, principles, or information. The problem of not discovering the correct information is becoming a larger issue as more and more information becomes accessible to any person through the internet, social media, etc. Also, the information that an individual might encounter could be wrong, and the trainee would have no guidance or feedback telling him or her that the information is not correct. Again, this is becoming more problematic with all of the false information available in today's world. Though in exploratory learning interventions individuals are actively engaged in constructing their own learning, they may not learn the important information without any guidance or may learn information that is not correct.

Exploratory learning is similar to error management training because both techniques give individuals little guidance and ask them to complete a goal. The main difference is that error management training places greater emphasis on the commission of errors and subsequently learning from those errors (Keith & Frese, 2005) whereas exploratory learning interventions are focused mostly on individuals discovering and learning information. When a person engages in exploratory learning or guided exploration, the person will commit many errors during learning. These errors in essence will surprise the learner, and he or she will devote more attention and cognitive resources to understanding and overcoming the errors (Bell & Kozlowski, 2010). Because of this, exploratory learning is often a common component of error management training.

Summary. In this study, I utilized PBL, TBL, and PI among the well-developed active learning techniques to examine my hypotheses. I used these three techniques because they are well-developed, widely utilized, and used in the population that participated in my study. Moreover, I examined the effects of the key components on outcomes by examining the effects of three active learning interventions on outcomes. I posited that the presence of the key components differs in the three interventions as indicated in Table 1.

Table 1

Hypothesized Variability of Key Components in the Different Active Learning Interventions

	Accountability	Feedback	Guided Exploration
PBL	Medium	Low	Low
TBL	High	High	High
PI	Low	Medium	High

Note. The PBL hypotheses are based on my classroom observations and prior research (e.g., Savery, 2006). The TBL hypotheses are based on my classroom observations and prior research (e.g., Michaelsen, Sweet, & Parmelee, 2011). The PI hypotheses are based on my classroom observations and prior research (e.g., Crouch & Mazur, 2001).

Real World Issues

Because this study was a field experiment conducted in a medical school setting, there were many issues and confounds that I could not control. I will address some of these issues and discuss how I tried to minimize their impact. One issue concerned the participant population pool. The total number of possible participants was 109 students. According to power analysis (*d* = 0.3), I needed a sample size of at least 134 participants to conduct bivariate correlations. Therefore, I did not have a sample with an adequate size to meet these requirements. Unfortunately, there was nothing that I could do to correct this issue.

Also, the nature of the curriculum of the school where this study was conducted introduced further confounds. First, the three active learning interventions were not administered separately. Second, participants were in different courses of different lengths with different instructors that focused on different content throughout the year. To overcome this issue I isolated pharmacology exam questions from each exam in order to measure learning. I did this because pharmacology was a consistent topic that was included in each course throughout the year. Third, though active learning techniques were the main method of content delivery during each course, participants had additional learning opportunities, which could have affected relationships with outcomes. Fourth, though most of the professors were trained concerning how to administer each active learning intervention, the professors invariably had different delivery styles such that all active learning interventions were not administered uniformly across courses. Fifth, TBL had been utilized in this medical school for over ten years whereas PI had been utilized for about three years and PBL has been utilized for about two years. Because of this, the faculty and students had more familiarity with TBL compared to PI and PBL. Finally, PBL was the final active learning intervention that was introduced to the participants during the academic year. However, when PBL was introduced most of the participants had shifted the focus of their studies from their classroom work to Step 1 of the United States Medical Licensure Exam. This exam was critical to the participants' future and many participants spent months preparing for it.

Because of this shift in focus, the introduction of PBL was not well received among many participants. I further discuss these issues and confounds in the method and discussion sections of this paper.

Method

Sample and Design

Participants (*N* = 103) were medical school students from a large public university in the Midwest. Students participated in exchange for the chance to win prizes (e.g., gift cards).

I used a quasi-experimental, repeated measures design without random assignment. I used naturally occurring samples, i.e., medical students enrolled in identified classes. There was one repeated factor, the active learning condition (i.e., TBL, PBL, and PI). Although this study included the same participants throughout the study, team membership was different in each of the active learning conditions. As a point of clarification, the institution where I conducted this study referred to the PBL intervention as “WrightQ.”

The medical school divided the academic year into different courses. Each student participated in each of the three active learning interventions over the course of the year. They participated in a course that included PI, at a different time during the year they participated in a course that included PI and TBL, and at a different time during the year they participated in a course that included PI, TBL, and PBL.

Active Learning Conditions

In this section I specify each of the active learning conditions I had access to in this study.

TBL. At the beginning of the academic year, students were assigned to a TBL team. There were six to eight people per team. The teams remained together through the entire academic year. Prior to each TBL session, students were assigned pre-class readings based on the information covered during that class session. TBL classes were held in a large lecture hall where all TBL teams could be together at the same time. At the beginning of each TBL session, students completed an individual assessment (usually multiple choice questions which they completed and turned in) related to their pre-class readings. Next, students worked with other members of their assigned TBL team to complete the identical assessment as a group. Students needed to agree on the answers as a team. They received immediate feedback on whether their answers were correct. Next, there was a short appeals process. Finally, for the remainder of the class session the professor engaged the students using discussion, lecture, and asked students or teams questions that they could answer using clickers. This process was repeated for each TBL class session that was conducted.

PBL. At the beginning of a course, students were assigned to PBL teams. These teams were different from their TBL teams. Again, there were between six and eight students per team. The PBL teams remained together for the entirety of the course. Unlike TBL, there were no assigned prior readings in the PBL condition. The teams did not meet all together in a large lecture hall. Rather, each team met in a separate, smaller room. Each team was assigned a facilitator who remained with the team for the remainder of the course. At the beginning of a PBL session, students participated in a “check-in” process. During the check-in, each individual commented about whatever they wanted to share with the group (e.g., fun things they did during the weekend, recent vacations). After the check-in, each student received a copy of an unfolding “case.” Each case described a scenario that a doctor might encounter regarding a single patient. Each case was divided into paragraphs, and each paragraph built upon the information included in the previous paragraph. Students took turns reading the paragraphs. After each paragraph, the team discussed the content it contained. Students were instructed to not read ahead. In their discussion of each paragraph, students discussed what they learned about the patient and gaps in the student's knowledge that they needed to fill in order to help the patient. This process was repeated for each paragraph. Once students completed their discussion of each part of the unfolding case, they summarized their discussions about the gaps in their knowledge and decided which topics they need to learn about before they came together again as a team for the next PBL session. During a following "check-out" process the students discussed the PBL session (e.g., group dynamics, whether they got off topic as a group, positive/negative feedback to individuals in the group). The team then dispersed, and each student studied the topics the group decided upon. Typically, each student studied every topic and was well-prepared to discuss each topic in the next session.

After a set amount of time (usually a few days), the students reconvened with their teams. Then, they repeated the check-in process and discussed all of the relevant content that they learned during their individual study time. The format of this portion of a PBL session involved discussion. Students could choose to comment as much or as little as they preferred. After the students discussed all of the topics that they agreed upon in the previous session, they completed a check-out process and the session ended.

Prior to each session, three members of the group were assigned specific roles. The students held these roles for one PBL session, and the roles rotated to other members of the team in the next PBL session. The roles were leader, scribe, and time-keeper. The leader was the person in charge of keeping order, ensuring the group is on topic, ensuring everything is covered that needs to be covered, etc. The scribe was in charge of taking notes and distributing those notes to the group after the session is over. The time-keeper was in charge of making sure the group does not spend too much time on one topic and helped the leader ensure that there was adequate time to discuss all the information that needed to be discussed.

A facilitator, typically a medical school professor, was present for each PBL session. The facilitator was not actively involved in any of the group discussion. Rather, he or she took part in the check-in and check-out processes but generally remained silent throughout the rest of the session. The facilitator was responsible for assisting the team in keeping order though the teams usually did not require much facilitator intervention. The facilitators would only interject into the discussion when he or she deemed it necessary, asking a probing question, getting the group back on topic, briefly clarifying points that the students do not understand, etc. One purpose was to teach students how to lead and participate in these types of teams, so facilitators would not interject too often in order to help students build these skills.

PI. For each PI session, students were assigned pre-readings before class. For each class period, students entered the classroom and were randomly assigned seats. The professor alternated between short lectures and multiple choice questions. If a certain percentage of students did not answer a multiple choice question correctly, then students were instructed to discuss that question with peers who were seated close to them. After a few minutes of peer discussion, students retook the same multiple choice question. The expectation was that a higher percentage of students would answer the question correctly the second time.

Structurally, PBL is much different from TBL and PI. One major difference is that PBL is conducted exclusively in small group settings in individual rooms whereas both TBL and PI are conducted in large lecture halls. Also, in PBL, participants only work through one case, and there is no feedback concerning right or wrong answers.

Key Components

Accountability. I measured accountability using the scale developed by Thomas, Dose, and Scott (2002, $\alpha = .86$). This measure consisted of nine items with responses on a 5-point scale, ranging from *Not at all* (1) to *Great Extent* (5). Scores were averaged, and higher averages indicate a higher level of accountability. An example item is "To what extent are your classmates interested in how well you perform your schoolwork?" In addition to collecting data from students, I enlisted two members of the medical school administration to rate accountability levels. These subject matter experts are highly qualified for this task. One of the raters earned an M.D. and is currently the Assistant Dean for Curriculum and Medical Education Research. This person has over 20 years of experience in medical education and research. The other rater earned a M.S. in leadership development and is currently the Manager of Curriculum and Instructional Design. Both of these individuals are prominently involved in the design and implementation of the curriculum at the medical school where they are employed. Thus, I had two forms of data relating to accountability: participants' perceptions of the levels of accountability and subject matter experts' assessments of the levels of accountability. The data obtained from the subject matter experts functioned as objective ratings and the data obtained from the students was intended to be used as a manipulation check. The complete measure is in Appendix A.

Feedback. I measured feedback using the scale developed by Steelman, Levy, and Snell (2004, $\alpha = .92$). This measure consisted of five items with responses on a 7-point scale, ranging from *Strongly disagree* (1) to *Strongly agree* (7). Items scores were averaged, and a higher average indicated a higher level of feedback quality. An example item is "My classmates (instructor) give me useful feedback about my classroom performance." In addition to collecting data from students, I enlisted two medical school faculty members to rate feedback quality levels. Thus, I had two assessments of data relating to feedback quality. The data obtained from the subject matter experts functioned as objective ratings and the data obtained from the students was intended to be used as a manipulation check. The complete measure is in Appendix B.

Guided exploration. A measure of guided exploration was created for this study. This measure consists of three items with responses on a 7-point scale, ranging from *Strongly disagree* (1) to *Strongly agree* (7). Scores were averaged, and a higher average indicated a higher level of guided exploration. An example item is "Rate the degree that a student's exploration is guided in the ___ course." Two medical school faculty members rated the degree to which the student's exploration is guided. The data obtained from the subject matter experts functioned as objective ratings and the data obtained from the students was intended to be used as a manipulation check. The complete measure is in Appendix C.

Demographics. I measured demographic information, including age, gender, and race. The complete measure is in Appendix D.

Primary Outcomes

Learning. Learning was assessed using the academic records of the students. I examined exam scores. I enlisted a member of the medical school faculty to determine which exam questions related to pharmacology in each of the courses. I did this because though each course focused on a different area of medicine (e.g., Medical Neuroscience, Reproduction), all of the courses included elements related to pharmacology. Therefore, in order to equate content across courses, I used only the questions that related to pharmacology on each as exam as my measure of exam scores.

Metacognition. I measured metacognition using one subsection of the measure developed by Wells and Cartwright-Hatton (2004, α = .92). This measure consisted of six items with responses on a 4-point scale, ranging from *Do not agree* (1) to *Agree very much* (4). Scores were averaged, and a higher average indicated a higher level of metacognition. An example item is “I pay close attention to the way my mind works.” The complete measure is in Appendix E.

Self-efficacy (state). I measured state specific self-efficacy using the scale developed by Riggs, Warka, Babasa, Betancourt, and Hooker (1994, α = .89). This measure consisted of ten items with responses on a 7-point scale, ranging from *Strongly disagree* (1) to *Strongly agree* (7). Scores were averaged, and a higher average indicated a higher level of state specific self-efficacy. An example item is “I have confidence in my ability to do well in ______ class.” The complete measure is in Appendix F.

Professionalism. I measured professionalism using a revised version of the scale developed by Chisholm, Cobb, Duke, McDuffie, and Kennedy (2006, α = .82). This measure consisted of six items with responses on a 7-point scale, ranging from *Strongly disagree* (1) to *Strongly agree* (7). Scores were averaged, and a higher average indicated a higher level of professionalism. An example item is “I follow through with my responsibilities.” The complete measure is in Appendix G.

Emotion Control. I measured emotion control using the scale developed by Keith (2005, α = .80). This measure consists of eight items with responses on a 5-point scale, ranging from *Strongly disagree* (1) to *Strongly agree* (5). Scores were averaged, and a higher average indicated a higher level of emotion control. An example item is “I was able to motivate myself to continue.” The complete measure is in Appendix H.

Qualitative Data

Qualitative data can offer a high quality explanation of results. As previously discussed, in order to obtain a richer explanation for my data and another form of high quality data, I decided to collect qualitative data in the form of participant interviews after the participants had participated all of the interventions. This would aid in the interpretation of my results. Qualitative data provided a more in-depth measure of how participants thought about the three active learning interventions. To obtain additional data and understanding of the three active learning interventions, I conducted one on one student interviews. Participants were invited to participate through an email to the entire class, and six self-selected students were the participants. The interviews lasted between forty five minutes and one hour. Because I had spent so much time sitting in the participants' classes and interacting with them, I had gained the participants' trust. Therefore, the participants were more likely to be willing to be open and honest with me in their conversations. During the interviews, approximately one third of the time was spent talking about each active learning intervention. Typical questions that were asked by the interviewer were: "What did you do for the peer instruction?" "What is team-based learning?" or "What are you doing in between WrightQ [PBL] sessions?" Each interviewee was asked similar questions. The order of discussion about active learning interventions was counterbalanced across all six participants.

After conducting the interviews, I transcribed the conversations. Then, using the Linguistic Inquiry Word Count Software (Tausczik & Pennebaker, 2010), I analyzed the language used by the interview participants. The Linguistic Inquiry Word Count Software includes predefined categories that contain certain words that the creators of the program assigned to the categories. For example, the category of time includes words such as age, start, usually, and later. After I uploaded the transcribed conversations into the program, the Linguistic Inquiry Work Count Software reported the proportion of words that were used that had been assigned to each category. The proportion was based on the number of words used per one thousand words. For example, one student used words that had been assigned to the category of time at a proportion of 7.95. That is, of each thousand words spoken, this student used 7.95 words relating to time.

Procedure

During the course that included PI, I sent students at the midpoint of the course online surveys that measure accountability, feedback, and guided exploration. At the end of the course, I sent students surveys that assess metacognition, task specific self-efficacy, professionalism, emotion control, and metacognition. I repeated this procedure of sending surveys at the midpoint and end of the course with the same surveys for the courses that include TBL and PBL. Table 2 describes the temporal sequence of the survey distribution. I distributed surveys during the Medical Neuroscience, Musculoskeletal/Integument, and Staying Alive courses. Each time I administered a survey, I attended the participants' class the day of the survey administration and asked the participants to take the survey. I had previously attended many of the classes as part of my participant observation. Then, I emailed the participants a link to the survey.   Also, I collected qualitative data in the form of interviews. These interviews were conducted after last survey was administered.

Table 2

Temporal Sequence of Survey Distribution

Date	Sept. – Oct.	Jan. - Jan.	Feb. - Feb.
Subject	Medical Neuroscience	Reproductive	Musculoskelatal/ Integument
Active Learning Intervention	PI	PI TBL	PI TBL 1 PBL

I recruited students via email using addresses provided by the medical school. Participants completed an informed consent process prior to completing any surveys. During this informed consent process, I asked students whether they were willing to allow access to their academic records. Participating students entered their names in order for me to be able to match their records with their survey responses.

With assistance from the medical school, I incentivized students to participants by offering them gift cards to various local restaurants. After each survey administration, participants who completed the surveys at that time point were entered in a random drawing to win a gift card. Three gift cards were rewarded per drawing. If the participant did not complete the surveys at that time point, he or she was not entered into that drawing. Thus, participants had the opportunity to receive gift cards totaling approximately $30 in value. Furthermore, if a student participated in all the surveys at the six time points, he or she was entered into a drawing to win one of three Fitbits or one of two iPads.

Results

Data Cleaning

One hundred and three medical students participated in this study. However, not every participant completed each survey at each time point. In total, eighty seven participants completed all the surveys at all of the six time points. One hundred participants completed the surveys at the first time point, and ninety seven, ninety eight, ninety eight, ninety seven, and ninety participants completed the surveys at the subsequent five time points, respectively. All participants who began a survey completed that survey. Therefore, no participant data was deleted due to missing data.

Descriptive Statistics

Of the original one hundred and three participants, fifty three were female and fifty were male. The average age of the participants was 24.56 years (*SD* = 2.43). Sixty one percent of the participants self-identified as White, fourteen percent self-identified as African American, and twelve percent self-identified as Asian.

Before I conducted any analyses concerning the three key components, I needed to decide whether to test my hypotheses using the data obtained from the medical school students or the data obtained from the subject matter experts. The data I obtained from the participants was a measure of the participants' perceptions of the key components whereas the data I obtained from the subject matter experts was intended to be a more objective measure of the key components. Ultimately, I chose to use the data I collected from the students for two reasons. First, there was little agreement between the two subject matter experts, which rendered any analyses very difficult. Second, the pattern of results of the data that I obtained from the more experienced of the two subject matter experts aligned with the pattern of results from the data that I obtained from the students. Therefore, I decided to use the student data in my analyses. In the paragraphs below, any data that is reported is data that I obtained from the medical school students.

Measure Evaluation

As previously mentioned, the measure of professionalism was a substantially revised version of the existing measure developed by Chisholm, Cobb, Duke, McDuffie, and Kennedy (2006). The measure of guided exploration was developed specifically for this study. Both of these measures was administered during three different courses. I conducted exploratory factor analyses on the professionalism and guided exploration measures to examine whether each reflected a single dimension as expected. For each measure separately, I first examined one-factor solutions. For the one-factor solutions, I examined the presence of substantial factor loadings for each measure's items. I defined a substantial loading on a factor as any loading above .3.

Professionalism. As mentioned above, the professionalism measure consisted of six items taken from the professionalism measure developed by Chisholm, Cobb, Duke, McDuffie, and Kennedy (2006). I revised this version by selecting six items from it. This measure was intended to reflect one dimension. To examine the psychometric properties of this measure, I conducted an exploratory factor analysis.

I examined the scree plots for the surveys that were administered during the three courses, which each provided evidence of one factor (See Figures 1, 2, and 3). The Cronbach's Alphas were .73 for Course 1, .62 for Course 2, and .88 for Course 3. Though the internal consistency reliability at Course 2 was lower than desirable, the internal consistency reliability values for Course 1 and 3 were acceptable. Analysis for each survey administration indicated that deleting items would not significantly improve the Cronbach's Alpha. Therefore, I did not delete any items.

Next, I conducted an exploratory factor analysis with a one-factor solution. For the one factor solution, factor loadings ranged from .37 to .77 at Course 1. At Course 2, the factor loadings ranged from .33 to .64. At Course 3, factor loadings ranged from .59 to .88. These loadings provided evidence of a one-factor solution for this measure (See Table 3).

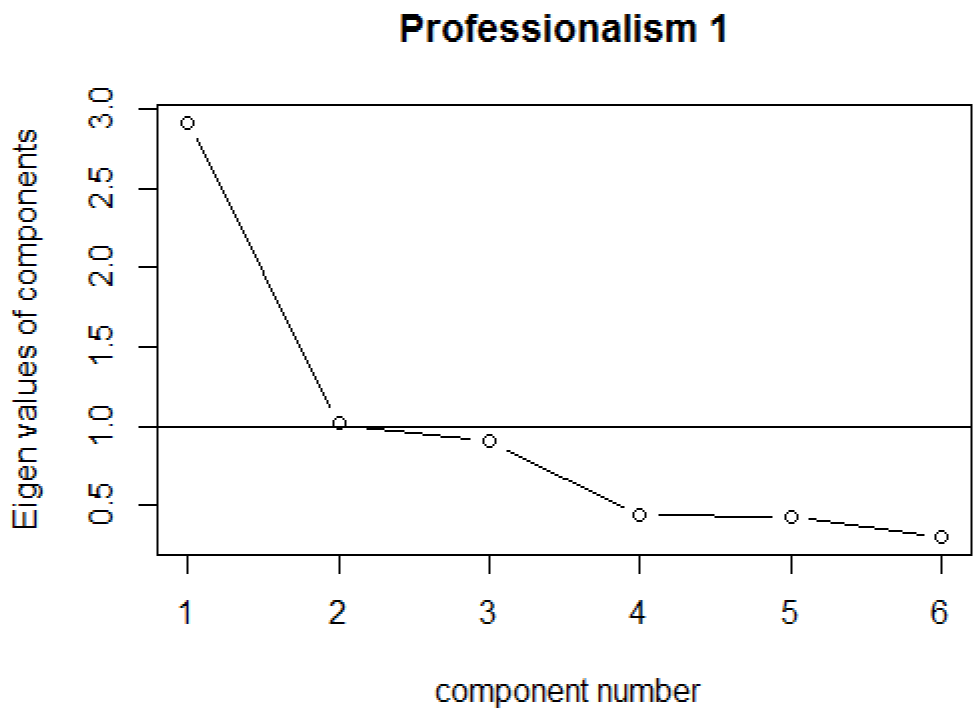


Figure 2. Scree plot of the professionalism measure for Course 1.

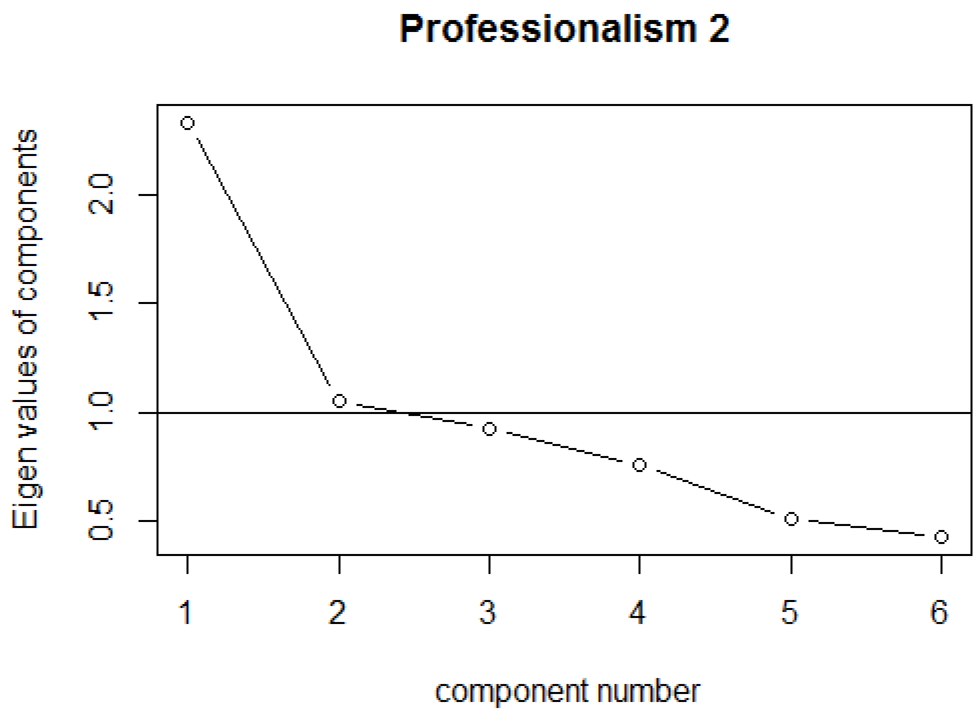


Figure 3. Scree plot of the professionalism measure for Course 2.

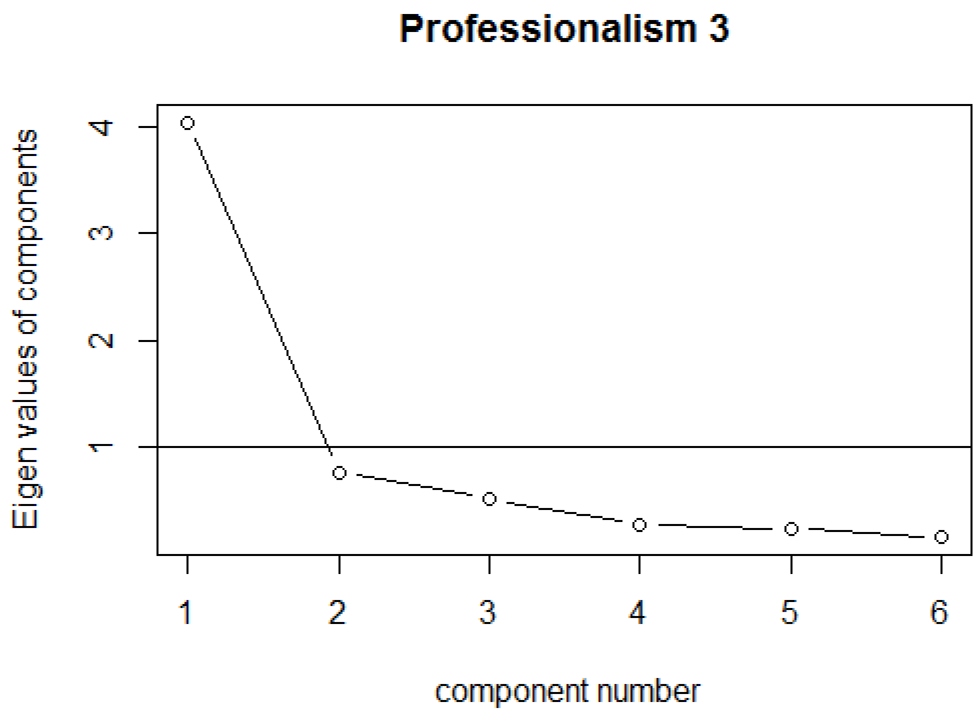


Figure 4. Scree plot of the professionalism measure for Course 3.

Table 3

Factor Loadings for a One-factor Solution for the Professionalism Measure for All Courses

Items	Course 1	Course 2	Course 3
If…I will be late, I contact the appropriate individual…	0.44	0.33	0.59
I accept consequences for my actions	0.77	0.44	0.86
I follow through with my responsibilities	0.76	0.60	0.88
I am committed to helping others	0.76	0.62	0.84
I would report a medication error…no one else was aware	0.54	0.44	0.76
I address others using appropriate names and titles	0.37	0.64	0.74

Guided exploration. The guided exploration measure was intended to reflect one dimension. The items were created specifically for this study. To examine the psychometric properties of the survey, I conducted an exploratory factor analysis for the survey given during each course.

First, I examined the scree plots which each provided evidence of one factor (See Figures 5, 6, and 7). The Cronbach's Alpha was .37 for Course 1. This alpha level was unacceptable. Due to this low alpha level, I changed two of the items in the guided exploration scale for the next two administrations of this measure. Specifically, I deleted the items "How often does the instructor offer suggestions to you regarding the process that you should use to obtain information relevant to the course" and "How often do you learn information related to the course without any suggestions or guidance from the instructor". I chose to delete these two items because they did not load onto the first factor when I conducted an exploratory factor analysis (See Table 4). I replaced those two items with two items that I thought would better measure the construct of guided exploration. The two new items were "How often do you have flexibility concerning what you can study outside of the classroom" and "How often are you given specific instructions about what to study in your ___ courses". The Cronbach's Alphas for the revised measure were .06 for Course 2 and .38 for Course 3. These alpha levels were unacceptable.

In an effort to better understand these results, I conducted an exploratory factor analysis with a one-factor solution. The factor loadings ranged from -.12 to .92 for Course 1. The factor loadings ranged from .02 to 1.00 at Course 3. The factor loadings ranged from -.28 to .61 at Course 3 (See Table 4).

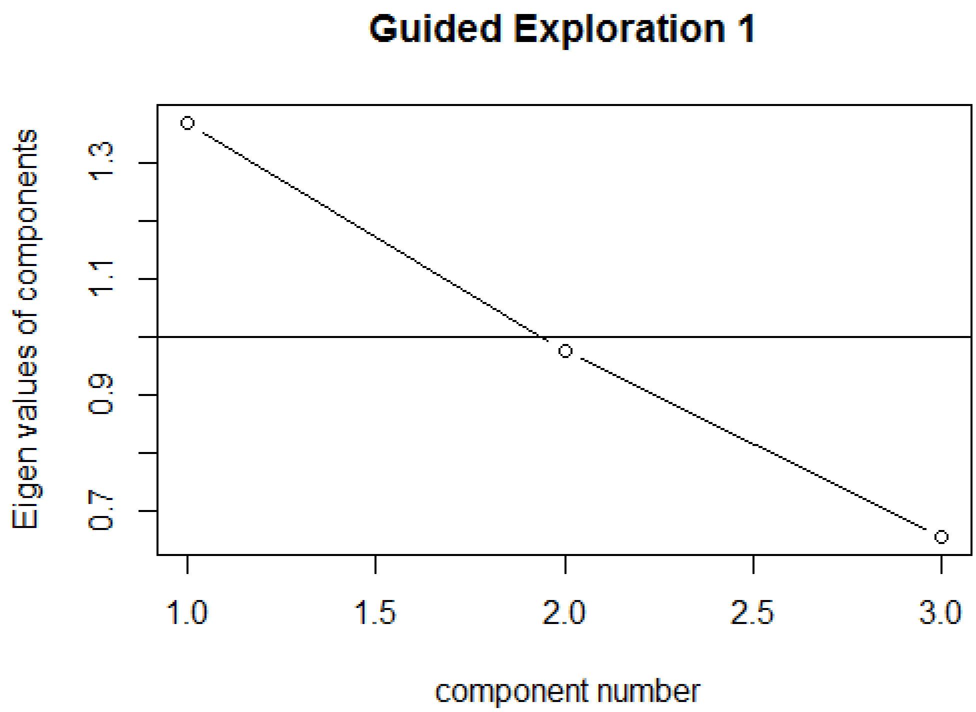


Figure 5. Scree plot of the guided exploration measure for Course 1.

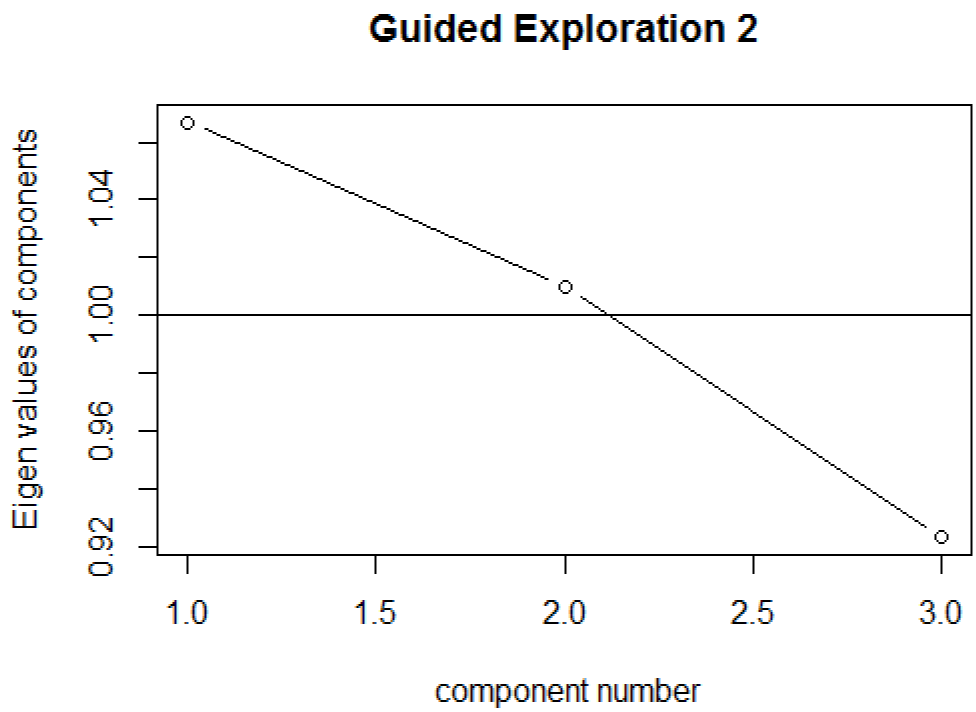


Figure 6. Scree plot of the guided exploration measure for Course 2.

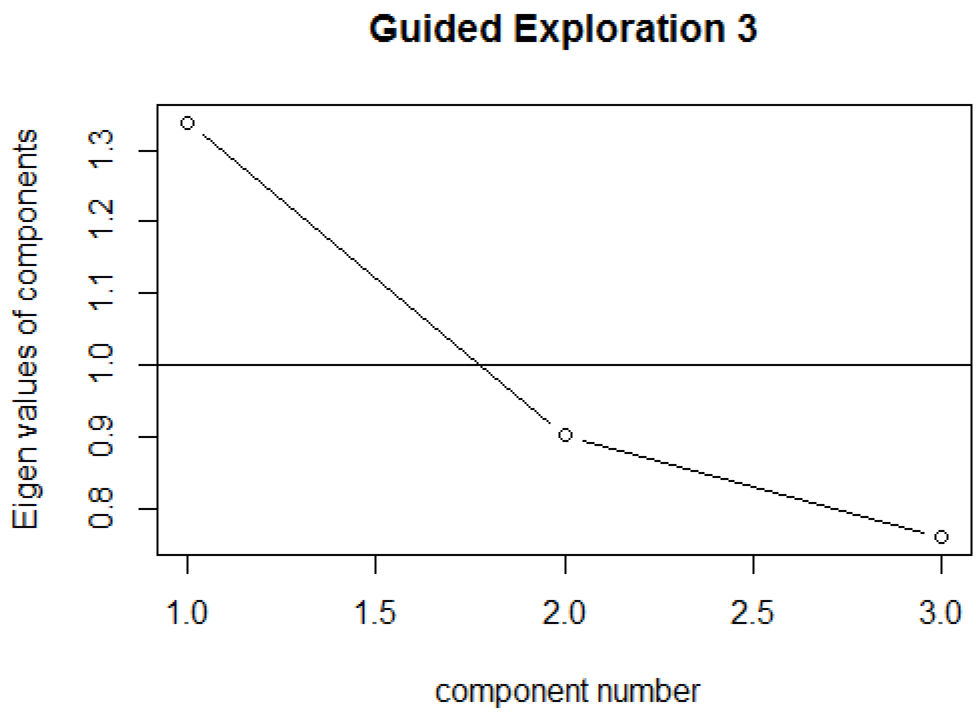


Figure 7. Scree plot of the guided exploration measure for Course 3.

Table 4

Factor Loading for a One-factor Solution for the Guided Exploration Measure for All Courses

Items	Course 1	Course 2	Course 3
How often does…instructor offer suggestions...examine…	0.92	1.00	-0.28
How often do you have flexibility concerning …study…	0.37	0.07	0.37
How often are you given specific instructions…to study…	-0.12	0.02	0.61

Based on these analyses, I concluded that the guided exploration measure did not constitute a valid measure of guided exploration. Therefore, I used the single item “How often does the instructor offer suggestions to you regarding the type of content you should examine outside of class" as my measure of guided exploration. I chose this item because it was the item that I used during all three survey administrations when I assessed levels of guided exploration. Also, based on the exploratory factor analyses, this item loaded onto the first factor during the survey administrations as part of Course1 and Course 2. Therefore, I concluded that that this item was the best measure of levels guided exploration compared to the other four items. Hereafter, any references to guided exploration refers to the above-identified single item.

Hypothesis Testing

Tables 5, 6, and 7 provide the means, standard deviations, and intercorrelations between measures used in subsequent analyses. Hypothesis 1 stated that feedback, accountability, and guided exploration would have positive relationships learning. I used end of course exam scores in each course as the criteria for learning. I tested Hypothesis 1 by examining the bivariate correlations between accountability, feedback, and guided exploration and learning. I examined the correlations for all three courses. For Course 1 (peer instruction), there was no significant relationship between feedback and learning ($r = 0.03$, $p > .05$), between accountability and learning ($r = 0.07$, $p > .05$), or between guided exploration and learning ($r = 0.02$, $p > .05$). For Course 2 (TBL), there was a significant relationship between feedback and learning ($r = 0.24$, $p < .05$), but there was not a significant relationship between accountability and learning ($r = 0.20$, $p > .05$). There was a significant relationship between guided exploration and learning ($r = 0.27$, $p < .05$), For Course 3 (PBL), there was no significant relationship between feedback and learning ($r = -0.01$, $p > .05$), between accountability and learning ($r = -0.13$, $p > .05$), or between guided exploration and learning ($r = -0.07$, $p > .05$).

Thus, Hypothesis 1 was not supported.

Table 5

Intercorrelations for Hypothesis Tests

	Course	Learning	Self-Efficacy	Metacognition	Emotion Control	Professionalism
Feedback	1	0.03	0.10	0.10	0.17	0.25*
	2	0.24*	0.21*	0.10	0.18	0.19
	3	-0.01	0.15	0.17	0.10	0.22*
Accountability	1	0.07	0.12	0.18	0.37*	0.34*
	2	0.20	0.09	0.27*	0.23*	0.31*
	3	-0.13	0.02	0.13	0.11	0.03
Guided Exploration	1	0.02	0.18	-0.13	-0.01	0.00
	2	0.27*	0.09	-0.02	-0.01	0.03
	3	-0.07	-0.22*	0.02	-0.09	-0.13

Note. *p <. 05

Table 6

Intercorrelations and Descriptive Statistics for Independent Variables

	Mean	*SD*	1	2	3	4	5	6	7	8
1. PI Feedback	4.89	1.26								
2. TBL Feedback	5.05	1.24	0.39**							
3. PBL Feedback	4.38	1.32	0.29**	0.50**						
4. PI Accountability	4.13	1.14	0.39**	0.24**	0.32**					
5. TBL Accountability	4.45	0.99	0.18	0.48**	0.31**	0.44**				
6. PBL Accountability	4.03	0.99	0.08	0.39**	0.50**	0.44**	0.54**			
7. PI Guided Exploration	4.86	2.01	0.11	0.10	0.16	-0.06	0.00	0.08		
8. TBL Guided Exploration	3.98	1.82	0.02	0.21*	0.04	0.13	0.15	0.17	-0.05	
9. PBL Guided Exploration	3.59	1.88	0.13	0.02	0.00	0.17	0.14	0.01	-0.03	-0.05

Note. $^{*}p <. 05$; $^{**}p < .01$.

Table 7

Intercorrelations and Descriptive Statistics for Dependent Variables

	Mean	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. PI Self-efficacy	4.94	1.38														
2. TBL Self-efficacy	5.00	1.16	0.54**													
3. PBL Self-efficacy	5.06	0.96	0.63**	0.70**												
4. PI Metacognition	5.35	1.19	0.06	-0.03	0.06											
5. TBL Metacognition	5.27	1.34	0.29**	0.18	0.03	0.76**										
6. PBL Metacognition	5.17	1.23	0.28**	0.13	0.13	0.69**	0.78**									
7. PI Emotion Control	4.98	1.02	0.43**	0.43**	0.58**	0.13	0.06	0.20								
8. TBL Emotion Control	5.09	1.17	0.32**	0.60**	0.52**	0.23*	0.37**	0.27*	0.63**							
9. PBL Emotion Control	5.09	1.15	0.34**	0.51**	0.55**	0.17	0.12	0.22*	0.75**	0.77**						
10. PI Professionalism	6.15	0.63	0.21*	0.12	0.25*	0.34**	0.28**	0.30**	0.37**	0.36**	0.40**					
11. TBL Professionalism	6.09	0.86	0.35**	0.50**	0.30*	0.29*	0.47**	0.38**	0.39**	0.58**	0.42**	0.68**				
12. PBL Professionalism	6.15	0.81	0.23*	0.28**	0.34**	0.22*	0.21	0.40**	0.20	0.25*	0.34**	0.52**	0.71**			
13. PI Learning	92.26	8.34	0.26*	0.31**	0.14	-0.03	0.12	0.03	0.14	0.30*	0.12	0.08	0.03	0.04		
14. TBL Learning	72.22	11.41	0.27*	0.38**	0.24*	0.14	0.22	0.22	0.16	0.22	0.28*	0.09	0.02	0.05	0.27*	
15. PBL Learning	91.55	15.08	0.33**	0.24*	0.26*	0.12	0.16	0.08	0.03	-0.01	-0.25	-0.18	-0.04	-0.10	0.14	0.23*

Note. $*p <. 05$; $**p < .01$.

Hypothesis 2 stated that feedback, accountability, and guided exploration would have positive relationships with self-efficacy. I tested Hypothesis 2 by examining the bivariate correlations between accountability, feedback, and guided exploration, and self-efficacy. I examined the correlations for all three courses. For Course 1 (peer instruction), there was no significant relationship between feedback and self-efficacy ($r = 0.10$, $p > .05$), between accountability and self-efficacy ($r = 0.12$, $p > .05$), or between guided exploration and self-efficacy ($r = 0.18$, $p > .05$). For Course 2 (TBL), there was a significant relationship between feedback and self-efficacy ($r = 0.21$, $p < .05$), but there was not a significant relationship between accountability and self-efficacy ($r = 0.09$, $p > .05$) or between guided exploration and self-efficacy ($r = 0.09$, $p > .05$). For Course 3 (PBL), there was no significant relationship between feedback and self-efficacy ($r = 0.15$, $p > .05$) or between accountability and self-efficacy ($r = 0.02$, $p > .05$), but there was a significant negative relationship between guided exploration and self-efficacy ($r = -0.22$, $p < .05$).

Thus, Hypothesis 2 was supported for one of the nine tests.

Hypothesis 3 stated that feedback, accountability, and guided exploration would have positive relationships with metacognition. I tested Hypothesis 3 by examining the bivariate correlations between accountability, feedback, and guided exploration and metacognition. I examined the correlations for all three courses. For Course 1 (peer instruction), there was no significant relationship between feedback and metacognition ($r = 0.10$, $p > .05$), between accountability and metacognition ($r = 0.18$, $p > .05$), or between guided exploration and metacognition ($r = -0.13$, $p > .05$). For Course 2 (TBL), there was no significant relationship between feedback and metacognition ($r = 0.10$, $p > .05$). There was a significant relationship between accountability and metacognition ($r = 0.27$, $p < .05$) but not between guided exploration and metacognition ($r = -0.02$, $p > .05$). For Course 3 (PBL), there was no significant relationship between feedback and metacognition ($r = 0.17$, $p > .05$), between accountability and metacognition ($r = 0.13$, $p > .05$), or between guided exploration and metacognition ($r = 0.02$, $p > .05$).

Thus, Hypothesis 3 was not supported.

Hypothesis 4 stated that feedback, accountability, and guided exploration would have positive relationships emotion control. I tested Hypothesis 4 by examining the bivariate correlations between accountability, feedback, and guided exploration and emotion control. I examined the correlations for all three courses. For Course 1 (peer instruction), there was no significant relationship between feedback and emotion control ($r = 0.17$, $p > .05$). There was a significant relationship between accountability and emotion control ($r = 0.37$, $p < .05$) but not between guided exploration and emotion control ($r = -0.01$, $p > .05$). For Course 2 (TBL), there was no significant relationship between feedback and emotion control ($r = 0.18$, $p > .05$). There was a significant relationship between accountability and emotion control ($r = 0.23$, $p < .05$) but not between guided exploration and emotion control ($r = -0.01$, $p > .05$). For Course 3 (PBL), there was no significant relationship between feedback and emotion control ($r = 0.10$, $p > .05$), between accountability and emotion control ($r = 0.11$, $p > .05$), or between guided exploration and emotion control ($r = -0.09$, $p > .05$).

Thus, Hypothesis 4 was supported for two of the nine tests.

Research Question 1 addressed whether feedback, accountability, and guided exploration would have positive relationships professionalism. I tested Research Question 1 by examining the bivariate correlations between accountability, feedback, and guided exploration and professionalism. I examined the correlations for all three courses. For Course 1 (peer instruction), there was a significant relationship between feedback and professionalism (r = 0.25, p < .05) and between accountability and professionalism (r = 0.34, p < .05) but not between guided exploration and professionalism (r = -0.00, p > 05). For Course 2 (TBL), there was no significant relationship between feedback and professionalism (r = 0.19, p > .05). There was a significant relationship between accountability and professionalism (r = 0.31, p < .05) but not between guided exploration and professionalism (r = 0.03, p > .05). For Course 3 (PBL), there was a significant relationship between feedback and professionalism (r = 0 .22, p < .05) but not between accountability and professionalism (r = 0.03, p > .05) or between guided exploration and professionalism (r = -0.13, p > .05).

Thus, Research Question 1 was supported for five of the nine tests.

Growth Curve Analyses

To test for the relationship between time and the three key components and the five outcomes, I conducted multilevel growth curve analyses. This was necessary due to the longitudinal nature of the data. I was interested in discovering if time had an effect on the key components or outcome variables. I tested all variables following the model comparison approach suggested Bliese and Ployhart (2002). This approach uses random coefficient modeling to progressively build and evaluate more complex models. Once each model was estimated, I used likelihood tests to contrast alternative models (Bliese & Ployhart, 2002).

Procedure. First, I estimated the intraclass correlation coefficients. Next, I established a simple regression model that did not include any random effects. Next, I estimated a model that included random intercepts but not random slopes. Then, I conducted a likelihood ratio test to determine which of the first two models fit the data the best. Next, I estimated a model that allowed both random intercepts and random slopes. The next step was to compare the model that allowed both random slopes and intercepts with the model that the previous log likelihood ratio test determined fit the data the best. Next, it was necessary to conduct tests of autocorrelation. Autocorrelation is the idea that responses closer together in time may be more strongly related to each other than responses further apart in time (Bliese & Ployhart, 2002). Also, within person errors may not be independent because responses may become more or less variable over time (Bliese & Ployhart, 2002). Failing to account for autocorrelation can lead to underestimation of the standard errors and inflated *t* values (Bliese & Ployhart, 2002). Following the test of autocorrelation, I conducted a test of heteroscedasticity. Heteroscedasticity is the idea that residuals for different cases have different variances (Bliese & Ployhart, 2002). This violates an assumption in growth curve modeling. Failing to account for this can lead to incorrect estimates of standard errors and therefore to possibly incorrect inferences (Bliese & Ployhart, 2002).

Analyses. First, I estimated the intraclass correlation coefficients for each of the five outcomes. The intraclass correlation coefficient for self-efficacy was 0.62. The intraclass correlation coefficient for metacognition was 0.78. The intraclass correlation coefficient for emotion control was 0.75. The intraclass correlation coefficient for professionalism was 0.72. The intraclass correlation coefficient for learning was 0.39. These results provided evidence for the necessity of analyzing my data using multilevel growth curve modeling. Therefore, I continued with the procedure described above.

Feedback. Next, I tested the relationship between time and participants' perceptions of feedback. This test indicated a significant difference between the model that allowed intercepts to vary compared to the model that did not allow either the intercepts or slopes to vary, χ^2_{diff} (1) = 34.62, $p < .05$. There was no significant difference between the model that allowed both the intercepts and slopes to vary compared to the model that allowed for variation in the intercepts but not the slopes, χ^2_{diff} (2) = 2.36, $p > .05$. These tests indicated that the model that allowed for variation in the intercepts but not the slopes fit the data best. Time had a similar and significant negative relationship with feedback perceptions across subjects. In addition, there was variability in individuals' average perceptions of feedback (See Table 8). Finally, analyses showed that there were no problems of autocorrelation and heteroscedasticity.

Accountability. Next, I tested the relationship between time and participants' perceptions of accountability. This test indicated a significant difference between the model that allowed intercepts to vary compared to the model that did not allow either the intercepts or slopes to vary, χ^2_{diff} (1) = 49.90, $p < .05$. There was no significant difference between the model that allowed both the intercepts and slopes to vary compared to the model that allowed for variation in the intercepts but not the slopes, χ^2_{diff} (2) = 1.69, $p > .05$. These tests indicated that the model that allowed for variation in the intercepts but not the slopes fit the data best. Time did not have a significant relationship with accountability perceptions across subjects. In addition, there was variability in individuals' average perceptions of accountability (See Table 8). Finally, analyses confirmed the absence of autocorrelation and heteroscedasticity concerns.

Guided exploration. Then, I tested the relationship between time and participants' perceptions of guided exploration. This test indicated there was not a significant difference between the model that allowed intercepts to vary compared to the model that did not allow either the intercepts or slopes to vary, χ^2_{diff} (1) = 0.00, $p > .05$. There was no significant difference between the model that allowed both the intercepts and slopes to vary compared to the model that allowed for no variation in the intercepts or the slopes, χ^2_{diff} (3) = 0.00, $p > .05$. These tests indicated that the model that allowed for no variation in the intercepts or slopes fit the data best. Time had a similar and significant negative relationship with guided exploration perceptions across subjects (See Table 8). Finally, analyses confirmed the absence of autocorrelation and heteroscedasticity concerns.

Self-efficacy. Next, I tested the relationship between time and participants' levels of self-efficacy. This test indicated a significant difference between the model that allowed intercepts to vary compared to the model that did not allow either the intercepts or slopes to vary, χ^2_{diff} (1) = 92.78, $p < .05$. The model that allowed both the intercepts and slopes to vary did not converge. These tests indicated that the model that allowed for variation in the intercepts but not the slopes fit the data best. Time did not have a significant relationship with self-efficacy across subjects. However, there was variability in individuals' average perceptions of self-efficacy (See Table 8). Also, analyses showed that there was no problem with autocorrelation, but there was heteroscedasticity. I accounted for this by including a term for heteroscedasticity in the models reported.

Metacognition. Then, I tested the relationship between time and participants' levels of metacognition. This test indicated a significant difference between the model that allowed intercepts to vary compared to the model that did not allow either the intercepts or slopes to vary, χ^2_{diff} (1) = 166.24, $p < .05$. There was no significant difference between the model that allowed both the intercepts and slopes to vary compared to the model that allowed for variation in the intercepts but not the slopes, χ^2_{diff} (2) = 4.07, $p > .05$. These tests indicated that the model that allowed for variation in the intercepts but not the slopes fit the data best. Time did not have a significant relationship with metacognition across subjects. In addition, there was variability in individuals' average perceptions of metacognition (See Table 8). Also, analyses showed that there was no problem with heteroscedasticity, but there was autocorrelation. I accounted for this by including a term for autocorrelation in the models reported.

Emotion control. Next, I tested the relationship between time and participants' levels of emotion control. This test indicated a significant difference between the model that allowed intercepts to vary compared to the model that did not allow either the intercepts or slopes to vary, χ^2_{diff} (1) = 145.58, $p < .05$. The model that allowed both the intercepts and slopes to vary did not converge. These tests indicated that the model that allowed for variation in the intercepts but not the slopes fit the data best. Time did not have a significant relationship with emotion control across subjects. In addition, there was variability in individuals' average perceptions of emotion control (See Table 8). Finally, analyses confirmed the absence of autocorrelation and heteroscedasticity concerns.

Professionalism. Next, I tested the relationship between time and participants' levels of professionalism. This test indicated a significant difference between the model that allowed intercepts to vary compared to the model that did not allow either the intercepts or slopes to vary, χ^2_{diff} (1) = 107.55, $p < .05$. There was a significant difference between the model that allowed both the intercepts and slopes to vary compared to the model that allowed for variation in the intercepts but not the slopes, χ^2_{diff} (2) = 22.76, $p < .05$. These tests indicated that the model that allowed for variation in both the intercepts and the slopes fit the data best. Time did not have a significant relationship with professionalism across subjects. In addition, there was variability in individuals' average perceptions of professionalism (See Table 8). Finally, analyses confirmed the absence of autocorrelation and heteroscedasticity concerns.

Learning. Finally, I tested the relationship between time and participants' exam scores. This test indicated a significant difference between the model that allowed intercepts to vary compared to the model that did not allow either the intercepts or slopes to vary, χ^2_{diff} (1) = 30.60, $p < .05$. The model that allowed both the intercepts and slopes to vary did not converge. These tests indicated that the model that allowed for variation in the intercepts but not the slopes fit the data best. Time had a similar and significant negative relationship with exam scores across subjects. In addition, there was variability in individuals' average exam scores (See Table 8). Finally, analyses showed that there was no problem with autocorrelation, but there was heteroscedasticity. I accounted for this by including a term for heteroscedasticity in the models reported.

Qualitative Data Analysis

My previous interactions with the participants likely encouraged them to be open and unafraid to express their opinions during the interviews. This led to high quality interview data. The results of the qualitative data analysis revealed that there were differences in how the participants as a group thought about the different active learning conditions. Specifically, results indicated differences in language in relation to four categories: insight, health, biology, and work.

Problem Based Learning

As discussed previously, PBL was structurally different that TBL and PI. The main differences related to the number of individuals in each intervention and the presence of correct answers. In the following sections, I discuss how participants spoke less about insight in their comments regarding PBL compared to their comments about TBL and PI. Furthermore, I discuss how participants spoke more about health and biology when describing PBL compared to their comments about TBL and PI. I use example words that were frequently used by the participants to highlight how the participants spoke about each category.

Table 8

Results from Growth Curve Analyses

Variable		Coefficient	*SE*	*Df*	*T*	*P*
Feedback	Intercept	5.03	0.12	190	40.95	< .05
	Time	-0.25	0.07	190	-3.46	< .05
Accountability	Intercept	4.24	0.10	190	41.93	< .05
	Time	-0.04	0.06	190	-0.65	> .05
Guided Exp.	Intercept	4.78	0.17	190	27.42	< .05
	Time	-0.64	0.13	190	-4.70	< .05
Self-efficacy	Intercept	4.70	0.12	188	40.26	< .05
	Time	0.01	0.05	188	0.24	> .05
Metacognition	Intercept	5.29	0.13	183	41.03	< .05
	Time	-0.06	0.05	183	-1.21	> .05
Emotion Con.	Intercept	4.98	0.11	183	44.16	< .05
	Time	0.05	0.04	183	1.23	> .05
Professionalism	Intercept	6.11	0.08	183	81.15	< .05
	Time	-0.03	0.04	183	-0.65	> .05
Learning	Intercept	90.58	0.98	147	92.05	< .05
	Time	-2.60	1.04	147	-2.50	< .05

Insight. The category of insight focused on words that pertain to obtaining information or becoming aware of something. This category included words such as discover, reveal, explain, and inform. Participants used fewer words related to insight when they spoke about

PBL compared to when they spoke about TBL and PI. The average number of spoken words that were related to insight per thousand words was 4.65 (95% CI: 4.27, 5.03) for PBL compared to 5.81 (95% CI: 5.11, 6.51) for TBL and 6.02 (95% CI: 5.12, 6.92) for PI. To illustrate this finding, I highlight the use of the words *answer* and *understand* below. Participants rarely used these two words when speaking about PBL. However, participants frequently used the word *answer* when talking about TBL. The following are two examples with respect to TBL.

> So the application isn't graded, so it's very just like your own benefit just listening and trying to *answer*.

> And there’s a group of six people in each TBL group, and we'll go through the same questions and try to discern which is the best *answer*.

Participants used the word *answer* in the following ways when they spoke about PI.

> You try to come to the right *answer*. Based on discussion…

> And then you click in your *answer* to what you think it is for the question.

Another key part of insight involves its source. In both PI and TBL, the source of the insight seemed to be the peers of the participants. For example, in TBL, participants spoke about how they gained understanding, sometimes pointing to the social aspects of the experience

> But after this discussion, it really helped me *understand* the gap. I guess I didn't catch during reading.

> That way we can kind of learn from each other and *understand* the thinking of the other people and hopefully come to the right answer.

> When participants spoke about PI, participants used the word *understand* in the following ways.

> Hopefully your friends, your peers, can help you solidify the concepts that you missed or you don't really *understand.*
>
> Oh, I discuss with as many people as possible. It depends on if I feel pretty confident in the answer, then I'm discussing, I'm confirming. If I feel like I know the answer, I'm more so confirming. If I don't know the answer, I'm more so trying to *understand.*

Health. The category of health focused on words that pertained to medicine and medical conditions. Words that were included in this category were: obese, pain, pill, and prescription. Participants used more words related to health when they talked about PBL compared to when they talked about PI and TBL. The average number of spoken words that were related to health per thousand words was 1.53 (95% CI: 1.22, 1.84) for PBL, compared to 0.63 (95% CI: 0.30, 0.96) for TBL and 0.51 (95% CI: 0.26, 0.75) for PI. To illustrate this point, I first highlight participants' use of the word *pathology* when they spoke about PBL.

> We go through each learning objective that we were previously assigned. So like the group leader will be like what's *pathology* of Aortic Stenosis?
>
> And then after that, we come up with learning objectives. About the disease that we just have diagnosed our patient with. Then like what's the *pathology* behind it was the physiology behind it.

Next I highlight the use of the word *pathophysiology*.

> Because then I can see everything that's going on, and then we put up the *pathophysiology* on the board.
>
> We struggled a little bit to come up with we could kind of through the *pathophysiology*. To come up with something that was happening.

Finally I highlight the use of the word *pharmacology*.

> I'm awful with *pharmacology*, but one of the guys in my group is like a walking drug list. I don't know how he remembers it all, but we'll bring up drugs, and he's like oh yeah this is the mechanism of action like this is what you know. So then I kind of get like a little mini lecture from him.

Biology. The category of biology related to words that focused on biological processes in the body. This category included words such as: muscle, nerve, organ, and cancer. Participants used words related to biology more frequently when they spoke about PBL compared to when they spoke about PI and TBL. The average number of spoken words that were related to biology per thousand words was 1.89 (95% CI: 1.52, 2.25) for PBL compared to 0.95 (95% CI: 0.56, 1.34) for TBL and 0.90 (95% CI: 0.46, 1.33) for PI. To highlight this point, below are quotes from participants when they talked about PBL. First, I focus on their use of the word *drug*.

> With peripheral vascular disease, so I actually took my whole group through each *drug* that you could prescribe in the mechanism of it. And the main side effects. What your timeframe window might be for something like a heart attack of when your enzymes are off. Which *drugs* are you going to use because of what's going on in the physio of it all to enhance it blocking whatever it might be for a *drug* and then if *drugs* aren't going to do it and you've got someone you know actively bleeding then you need a procedure and then what are your procedural options.

Next, I focus on their use of the word *heart*.

> I said if you had a *heart* attack in all three cases women present differently than men. And, then what if you have a twenty five year old versus a ninety five year

> old. That twenty five year olds having a *heart* attack probably for a way different reason.

Finally, I focus on the use of the word *diagnosis*.

> The biggest the biggest thing for yesterday was actually coming up with the differential *diagnosis* so that's always the first thing is they'll give us a very short thing, sometimes a sentence. And sometimes a few sentences. And the first question is always what's your different *diagnosis*.
>
> Start to discuss kind of a differential *diagnosis* based on just the history alone. Then you add in some more information whether that's the social history, family history, or physical exam. Again revisit kind of your differential *diagnosis*. Then you'll kind of theoretically order tests that you want. And so that's day one. And at the end of day one then once you kind of have a general idea of what your *diagnosis* is based on the test results based on the history all of that. Then, as a group you come up with your learning objectives and for us we kind of do the same ones every time we want to know. The pathophys we want to know the treatment and *diagnosis* the prognosis.

Preference

To conclude, I discuss participants' preferences. I do so because participants' preferences are instructive regarding how participants thought about and compared the active learning interventions to each other. At the end of five of the six interviews, the participants were asked to compare the three active learning interventions to each other. The other participant was not asked about his or her preference due to time constraints. Inevitably, the students mentioned which intervention they enjoyed most. Overall, four of the five participants explicitly stated that PI was their favorite of the three active learning interventions. The other participant stated that he or she could not decide between whether he or she liked PI or TBL more. None of the students stated that PBL was their favorite of the active learning interventions. The following quotes are participants' full responses.

Student 1

I: Do you enjoy peer instruction?

S: Yes, that's the one I think I like the most.

I: Why is that?

S: I really like it because it's really engaging to students. I couldn't stand just sitting in a lecture because it's really kind of like it's very passive make me sleepy. But at the same time it is not throwing a dart in the dark just like I am you know find a question is like WrightQ [PBL]. I really like P.I. because we have a certain set of objectives and reading assignments and we go through. And then during the discussion we have someone that's an expert that really explains to us things that we missed or things that's not in the book. So I think it is very helpful. Compared to you know for example WrightQ [PBL] where a facilitator is there's not supposed to help us and direct us going to the right concepts. So that's why I really like P.I. and plus with P.I if everyone understands he won't spend my time to lecture us whereas if people as a whole group are confused he would go in depth. With lecture there is no feedback he just explains everything including the things that we know. And explains things that we don't know and really quickly. I really like P.I. because I feel like it's very interactive between the facilitator and the students.

Student 2

I: Let’s compare WrightQ [PBL] and TBL.

S: I think. TBL was also helpful because you're getting in WrightQ [PBL] eight different minds thinking about the same learning objective. And sometimes we can all be on the same track but we’re missing something that we're supposed to know. And then in TBL you have the facilitator there who's been through this and who knows what you need to know about and so he's like over it you should probably know this. I think that’s what WrightQ [PBL] misses.

I: What about peer instruction? Is there a different reason why you would choose peer instruction over TBL?

S: Peer instruction is just like a rapid fire questioning, so it's like identifying what you know and don't know.

I: Let's compare peer instruction to TBL. Do you like one better than the other?

S: I like them both I couldn't choose one or the other. I think they're both beneficial in different ways. TBL is very helpful. I feel like you think a lot more because of the application part. And you also have to discuss it with your group. And the discussions are definitely more in-depth in TBL then they are in peer instruction. Because peer instruction is very like. It's more superficial it's like do you know what or do you don't. And if you happen to find someone who knows it and they can kind of explain it to you. But you're limited in the time you can have to discuss the question. So that's also kind of what hinders the discussion from going really in-depth. Whereas in TBL like you have. They generally don't like pressures to get in a certain time they let us discuss it to a full length. So it's really you get more in-depth discussion about the question.

Student 3

S: A well done peer instruction is worth more time than anything if you were to remove my lectures.

Student 4

S: I think that my favorite is peer instruction. Like I said earlier just because I'm doing fine with sitting on my own just because I'm having to. You're learning to you're learning what works for you and I think that that's been the most helpful to kind of gauge where I am. And if I am in a peer instruction we get through a lot of questions in a two hour time because it's very fast paced. So you're getting tested on a lot of material verses a WrightQ [PBL] where we're only looking at one particular disease. And it seems like a lot of time to be putting into one thing granted you know it pretty well by the end of that. But there's a lot else out there versus a peer instruction you're going through a lot and while peer instruction isn't all encompassing, it's just to me a kind of identifies I'm like oh I really don't know a lot about hypertension. I need to go back to look at hypertension versus like I got a lot of the questions right on lipid problems so I think I'm a least I should be or close where I should be with lipid disorder. So to me it's kind of like a really quick benchmark to see where I am. I don't think you could do to all peer instructions I think that you need some sort of quizzes in there and you need the case sessions in there because you need to be good at that but I don't know the perfect way to balance all of those.

Student 5

S: I love peer instruction. I loved having peer instruction basically every day because it kept me on my reading schedule. It showed me what I did and didn't know and it kept me on my toes. It's super easy to get lazy when I don't have something to prove every day. That being said, I think TBL is what really keeps you honest in terms of how much you know. You have to take that IRAT on your own. You have to stare that reality in the face when you're like shoot I have zero clue to even like how to eliminate one of these. Whereas in peer instruction everyone's a little bit more laid back if you don't meet 80 as a whole you talk to someone around you ok you said C like you know so it's good but it's easier to kind of lax there whereas I feel like a TBL at least once a week. You know definitely like I said keeps you honest.

Discussion

Study Purpose

Active learning techniques have positive effects (e.g., Prince, 2004), but research has yet to determine the key components that contribute to the positive effects associated with active learning. The purpose of this study was to examine key components of active learning techniques that I theorized would contribute to the positive relationships associated with active learning. I hypothesized that feedback, accountability, and guided exploration were three components associated with the positive relationships between active learning and outcomes. I failed to find substantial support for any of these hypotheses. However, I did find results that are worth discussing. First, participants perceived differences in the presence of levels of feedback and accountability that were consistent with theory in all three interventions but failed to perceive theorized differences in levels of guided exploration. Second, all three active learning components had weaker than expected relationships with outcomes although there were a greater number of observed relationships related to levels of accountability and feedback than to levels of guided exploration. Possible explanations for this included my inability to adequately measure levels of guided exploration.

Further, my results raised issues related to each of the three key components. First, the nature of the questions that were chosen, interactions between participants and peers, and the source of feedback might have influenced results concerning feedback. Second, team membership might have influenced results concerning perceptions of accountability. Third, measurement issues might have influenced results concerning guided exploration. Fourth, the nature of each active learning intervention, study confounds, and measurement issues might have influenced results. In the next sections, I discuss my hypothesis testing, implications, and future research pertaining to feedback, accountability, and guided exploration.

Hypothesis Testing

Contrary to my hypotheses, I found that feedback, accountability, and guided exploration had no substantial relationships with learning, self-efficacy, metacognition, emotion control, or professionalism. However, three observed results are worth discussing. First, feedback had a positive relationship with professionalism in both PI and PBL. Second, accountability had positive relationships with emotion control and professionalism in both PI and TBL. Third, I failed to observe a change in participants' perceptions of accountability, self-efficacy, emotion control, or professionalism over time although perceived feedback, perceived guided exploration, and learning declined over time. Furthermore, there were significant individual intercept differences across time. I have organized the following discussion by key component, addressing for each results, issues, and limitations. I follow that with a discussion of general issues, implications, and limitations.

Feedback

Results. Participants perceived differences in the predicted direction concerning feedback levels in each of the three active learning conditions. Not surprisingly, the more experienced subject matter expert perceived similar levels of feedback as the participants. However, the relationships between feedback and outcomes were weaker than expected. I found that four of the fifteen hypothesized relationships between feedback and outcomes were significant. Also, the growth curve models showed a similar pattern of weak relationships concerning feedback.

More specifically, I found evidence that supports the idea that feedback has a positive relationship with professionalism. Research has shown mixed effects concerning the relationship between feedback and professionalism (e.g., Brinkman et al., 2007). I found that in both PI and PBL settings, feedback had a positive relationship with professionalism. As individuals received feedback, they desired to improve performance and make changes. In this study, participants were asked about their perceptions of the feedback that they received from their peers. Due to the structure of both PI and PBL, participants frequently interacted with and received feedback from their peers. In many instances, participants received feedback concerning whether answers were correct or whether they understood a concept. As they received this feedback from their peers, participants were likely to self-assess their own actions toward their peers and identify areas in which they needed to improve. This could have led to a desire to make changes that could relate to acting in a professional manner toward their peers, such as following through with responsibilities or reporting errors. This finding provided evidence that an increase in the levels of feedback may lead to an increase in professionalism. However, more research is needed to justify this conclusion.

Caveats. Though I found weak relationships between feedback and outcomes, I found that participants perceived differences in the levels of feedback across the three active learning interventions that were consistent with theory. I found that participants perceived higher levels of feedback in PI and TBL compared to PBL. I believe three issues contributed to participants' perceptions of feedback. These issues related to the nature of the questions posed to participants, the nature of interactions between participants, and the source of feedback.

First, participants' perceptions of feedback seemed to covary based on the extent to which questions had correct or incorrect answers. Participants perceived higher levels of feedback in interventions that included questions that had definitive answers. The structure of both PI and TBL required that participants answer multiple choice questions that had a correct answer. In each PI and TBL session, participants received feedback from the instructor concerning whether the participants' answers were correct or not. This feedback occurred immediately and frequently throughout the class period. In contrast, as noted in earlier, in a PBL session, participants were never given any multiple choice quizzes and were rarely given feedback from their facilitator. If a facilitator asked the participants a question, a correct answer was rarely identified. Furthermore, many of the questions that participants answered in PBL were implicit, and therefore no correct answer was ever given to participants. The lack of a correct answer in PBL might be a reason why participants experienced less insight in PBL compared to TBL and PI. This interpretation is strengthened by the fact that the interviewees reported greater insight in both TBL and PI compared to PBL.

Second, the nature of interactions between participants might have influenced differences in perceptions of feedback. All three active learning interventions included substantial interaction between participants. However, there were differences in perceptions of feedback across interactions (PI $M = 4.89$; TBL $M = 5.05$; PBL $M = 4.38$). Because there was a high degree of interaction among participants in all three active learning interventions but high perceived levels of feedback in only PI and TBL, this led me to the conclusion that the amount of interaction among participants did not have much influence with the perception of feedback. Instead, what affected perceived levels of feedback might have been the content discussed among participants. In PI and TBL, participants generally discussed a question that had been posed to them that had a correct answer. Participants attempted to glean knowledge from each other in order to discover the correct answers to questions presented by the professor. In PBL, participants' interactions did not revolve around answers to questions or choosing answers to questions. This was because no questions were presented to participants. Instead, interactions revolved around discussing material that the group needed to learn and discussing material that individuals had learned previously.

Also, the idea that interactions between participants influenced differences in perceptions of feedback was strengthened by the fact that interviewees did not gain similar insight from their peers in a PBL session compared to PI and TBL sessions. This is a bit disconcerting due to the nature of PBL. A primary goal of any PBL intervention is for the group members to teach each other. Based on my analyses, it appears as if this goal was not attained. Therefore, if perceptions of feedback are an accurate representation of actual feedback levels, it might be reasonable to conclude that in order to increase levels of feedback, researchers and practioners should include not only group discussion but group discussion that involves the participants talking about a question with a correct answer.

Third, the source of feedback might have influenced differences in perceived feedback. In all three interventions, participants received feedback from their peers. However, in PI and TBL, participants received feedback also from their instructors. In PI and TBL, the instructors identified the correct answers to questions and answered additional questions that participants may have had. In contrast to TBL and PI, the facilitator in a PBL session was a more passive participant in the process, rarely giving input that could have been perceived as feedback from the perspective of a participant. In PBL, participants received feedback, e.g., with respect to correct answers, primarily from their teammates.

Limitations. There are multiple possible explanations for the small relationships observed between feedback and outcomes. One possible explanation is that I measured only one aspect of feedback. I measured participants' perceptions of feedback from their peers. There are many other aspects of feedback, including specificity, frequency, and timing, that I did not measure that could have affected relationships with outcomes. Another explanation is that feedback might not be an important key component that influenced outcomes in any of the active learning interventions. Active learning interventions incorporate many different components. Some of these possible components include emotion control strategies, error framing, time spent on the subject, decision making, repetition, and whether the intervention was graded. Future researchers should consider these variables and others when attempting to understand key components of active learning.

Accountability

Results. Participants perceived higher levels of accountability in TBL compared to PI and PBL Again, the relationships between levels of accountability and outcomes were weaker than expected. I found four out of fifteen possible significant relationships. Similarly, the growth curve models produced a pattern of weak results.

More, specifically, I found that accountability had a positive relationship with emotion control and professionalism in both PI and TBL settings. Accountability refers to an implicit or explicit expectation that a person might hold relating to the need to justify his or her actions to others (Giessner, Van Knippenberg, Van Ginkel, & Sleebos 2013; Scott & Lyman, 1968). Emotion control is the ability to keep negative emotions under control (Keith & Frese, 2005). It is difficult and sometimes embarrassing to have negative emotional outbreaks. Therefore, as the level of accountability increases, which increases the need to justify actions, the desire to control one's emotions might increase because the individual might not want to have to justify a negative emotional outbreak. Similarly, the reason for the positive relationship between accountability and professionalism could be because as an individual feels more accountable he or she might want more to avoid the need to justify unprofessional behaviors.

Issues. It is important to try to understand why participants perceived higher levels of accountability in TBL compared to PI and PBL. Perhaps, team membership was the main factor that contributed to perceptions of accountability. In relation to accountability, the three interventions raised issues related to the length of team membership, presence of team quizzes, presence of peer evaluations, presence of grades, and amount of work.

First, in TBL, participants were members of the same team for the entire academic year. Research has shown that time can play a role in team effectiveness (e.g., Harrison, Mohammed, McGrath, Florey, & Vanderstoep, 2003). Perhaps, the longer teams are together, the more accountable individuals feel toward their teammates. Over the course of the academic year, TBL teams spent time together almost weekly in TBL sessions.

Second, in TBL, team members took a graded quiz together each time they were in a TBL session. This team quiz counted toward the individual participant's final course grade. As high achieving medical school students, participants in the TBL teams desired to earn the best grades possible. This could have led each team member to not only study to achieve a high individual grade but to study in order to help the team earn the highest grade possible. Over the course of the year, the team members came to know and understand the strengths and weaknesses of each of their team members and incorporated that understanding into team quizzes. Based on the interviews that I conducted, there were few participants who did not feel accountable to be prepared and ready to assist the team during team quizzes in order to help the team earn the highest grade possible.

Third, each team member was evaluated by his or her peers concerning topics such as his or her preparedness for class and his or her contributions to the team. These peer evaluations held each participant accountable to his or her teammates, and part of each participant's grade was based on the peer evaluations. Again, as high achieving medical school students, participants desired to earn the highest grades possible. Therefore, it is possible that the peer evaluations increased levels of perceived accountability.

However, these three points lead to an interesting discussion concerning how grades might have influenced perceived accountability. Although three factors I previously mentioned could have influenced perceptions of accountability, these three factors were similar in TBL and PBL. In both TBL and PBL, participants were together in teams for an extended period of time.

TBL teams were together longer, the entire academic year, but PBL teams were together for almost ten weeks. Also, during the ten weeks that the PBL teams were together, they met twice a week almost every week. Furthermore, peer evaluations were part of the structure of PBL, and each participant knew that he or she would be evaluated by his or her peers, similar to TBL. This raises the question as to why participants did not perceive similar levels of accountability. One reason for this could be because PBL was not graded. Also, there were no team quizzes during PBL sessions. Due to lack of grading, it is possible that participants did not place the same emphasis on being prepared for PBL as they would have for a TBL session. Furthermore, participants might not have given as much attention to the levels of preparation of their teammates.

One of the main purposes of PBL was for the participants to learn material individually, and there was no incentive for the participants to be good teammates and contribute to the discussion. However, as indicated before, most participants came well-prepared to PBL classes, and though some participants were more actively engaged in the PBL process than others, most participants contributed to the discussion. Therefore, participants were engaged but did not feel accountable to their teammates. This is because participants in some sense perceived their teammates simply as sources of information. Because there was not a goal (i.e., team quiz) for the team to work toward together and the purpose of PBL was to individually master a topic, it is possible that participants in PBL did not feel accountable to their teammates because they had no reason to be held accountable.

Limitations. There could be many explanations for why I observed a pattern of relationships that was consistent with theory but the effects of accountability were weaker than I anticipated. First, as I discussed in the previous section, accountability might not be a key component of active learning that affected outcomes. Though accountability is a component that was present in all interventions, it might not have been an important influence on outcomes. Or in contrast, accountability could have been equally important in all interventions. Second, I might not have captured important aspects of accountability. I measured the level of accountability that participants felt toward their peers. However, there are other aspects of accountability that I did not measure. I did not measure perceptions of accountability that participants felt toward the instructor. That might have been a more powerful aspect of accountability than accountability to peers. Also, I did not measure how accountable participants felt to the medical school in which they were enrolled. This could have affected participant affect or buy-in, which could have had relationships with outcomes. Finally, I did not measure how accountable participants felt to themselves. Each of these aspects of accountability might have affected outcomes in different ways.

Guided Exploration

Results. Participants did not perceive differences in levels of guided exploration. Also, the subject matter experts did not agree concerning differences in levels of guided exploration. Based on my hypotheses, participants should have perceived lower levels of guided exploration in PBL compared to TBL and PI. However, I found only two out of fifteen possible significant relationships. Similarly, the growth curve models produced a pattern of weak results. Interestingly, the means for the guided exploration measures were higher for both PI and TBL compared to PBL, but the measures had large standard deviations. The standard deviations observed for the guided exploration measures were almost double the standard deviations of the feedback and accountability measures. The most likely reason for the large amount of variance was the difficulty that I had developing an adequate guided exploration measure.

Issues. There are many explanations for why I did not find the results that I expected. One explanation concerns the measurement of guided exploration. In this study, I found it difficult to measure guided exploration. No previous attempts had been made to measure levels of guided exploration with a survey measure. Previous researchers who examined guided exploration placed participants into groups of guided or non-guided interventions (e.g., Bell & Kozlowski, 2008) without any measure of guided exploration. However, I examined naturally occurring levels of guided exploration. I chose to do this because based on my research and previous knowledge of PI, TBL, and PBL, I believed that participants would experience different levels of guided exploration due to the nature of each intervention.

Given this, I believed the best way to measure guided exploration was to develop and administer a survey that captured guided exploration on a single dimension. Ideally, it would have been beneficial to develop and administer a measure of guided exploration that included more than three items. This would have allowed me to more thoroughly measure the construct of guided exploration. However, I decided to use three items to lessen the amount of time that the participants needed to spend taking each survey. As previously mentioned, I was very cognizant of time requirements for completing surveys because the participants were medical school students who have very little free time. Also, I administered surveys that were as short as possible to increase the sample size and retain as large a number of participants as possible. After I developed and administered the measure of guided exploration at Course 1, I conducted factor analytic work to determine the efficacy of the measure. The factor analyses showed that my first measure of guided exploration was not adequate. Therefore, I replaced two of the three items and administered the measure two more times. Then, I conducted more factor analyses, and the measure was still inadequate.

Future research. Given my difficulties developing a measure of guided exploration, one might explore alternatives to measuring naturally occurring levels of guided exploration. It might be more efficacious for researchers or practitioners to pre-determine if an intervention is high or low in levels of guided exploration, as has been done in previous research studies.

However, if a researcher is interested in administering a measure that captures guided exploration, I suggest the following. First, examine a larger initial item set. This is more consistent with accepted procedures concerning scale development (see Hinkin, 1995, for a more thorough review of scale development). Also, when assessing guided exploration, differentiate between exploration that occurs in a classroom settingand exploration that occurs outside of a classroom setting. Prime respondents to know that they will be asked questions about both settings. This is because an intervention could theoretically be high in guided exploration but not necessarily high in guided exploration in both settings. PBL would be an example in which guided exploration should be high both in and outside of the classroom. The setting of guided exploration could lead to differential effects on outcomes, which would be a potential interesting area of future research.

Second, the questions that I used as items for the guided exploration measure focused on how participants chose what to study. For example, one of the items was “How often are you given specific instructions about what to study for your PBL sessions?” However, the construct of guided exploration encompasses more than simply the study of material. Guided exploration encompasses how individuals use learning materials, think about and use errors, and engage in problem solving. A well-developed measure of guided exploration should include items that capture all aspects of guided exploration and not be as narrowly focused as the measure that I administered.

General Issues

Differential Effects of Key Components within TBL, PI, and PBL. I observed more significant relationships between key components and outcomes for tests that were conducted in a TBL setting compared to tests that were conducted in PBL and PI settings. I observed a predicted significant relationship in six of the possible fifteen hypotheses tests that were conducted in a TBL setting. However, I found only two out of fifteen tests that were conducted in a PBL setting to be significant. Furthermore, I found only three out of fifteen tests that were conducted in a PI setting were significant. An explanation for these results could be that because the perceived levels of feedback and accountability were higher in the TBL setting compared to PBL and PI settings (See Table 6), feedback and accountability had a greater influence on the variability within outcomes in a TBL setting. I believe there could be two possible explanations for the differences in the perceptions of feedback and accountability in TBL compared to PBL and PI.

First, I believe that the length of time that teams were together in TBL contributed to the results that I observed in TBL settings. As previously discussed, TBL teams were together for an entire academic year, often meeting multiple times per week. In PBL settings teams were together for a matter of weeks, and in PI there was no group dynamic that occurred for more than one class period. Because of the difference in the amount of time spent together, participants could have perceived that they received more feedback from their peers in TBL simply because there were more opportunities to receive feedback. Concerning perceptions of accountability, perhaps the longer a team is together the more accountable a team mate will feel to his or her team members. These reasons could have contributed to the higher levels of feedback and accountability that participants perceived in the TBL setting compared to the PBL and PI settings.

Second, I believe that the length of the class sessions could have contributed to these results. TBL sessions were usually longer than both PI and PBL sessions. Typical TBL sessions occurred for about two and one half hours, whereas, typical PBL and PI sessions lasted about one hour. Because of this discrepancy, participants had more opportunities to receive feedback or be held accountable in a single class period and over the course of the academic year. This could have contributed to the higher levels of feedback and accountability that participants perceived in the TBL setting compared to the PBL and PI settings.

Emotion control compared to metacognition. Another interesting result I observed in this study was the relationship between the key components and emotion control versus the relationship with metacognition. I hypothesized that the key components would have similar relationships with both outcomes. However, I observed significant relationships for three of the nine possible relationships in which emotion control was the outcome compared to only one of the nine possible significant relationships in which metacognition was the outcome.

One reason for these results might be that my measures of accountability and feedback focused on interactions between peers. These interactions could have led to greater affective changes (i.e., emotion control) than cognitive changes (i.e., metacognition). Interactions with peers could require more affective skills than cognitive skills relative to interactions with professors. This is due to the social nature of the interactions between peers in active learning interventions. It is possible that if my measures focused on interactions between the participants and faculty members, I might have observed stronger relationships with metacognition. Though active learning interventions should have positive relationships with both cognitive and affective outcomes, future researchers could examine whether active learning techniques have differential effects on cognitive and affective outcomes.

Individual learning compared to group learning. One interesting finding worth discussing concerns the nature of PBL compared to PI and TBL. Based upon the results that I found, participants potentially perceived PBL as a more individual learning activity instead of a group learning activity. All of the in-class activities related to PBL were done within teams and the same teams throughout the year. However, participants perceived PI as more of a group learning activity than they did PBL. This is interesting because there were no structured teams in a PI session, and participants were randomly assigned different seats for each class period. This meant that they had less of a prior relationship with their fellow participants in a PI session compared to a PBL session.

One reason why participants might view PBL as more of an individual learning activity was because there were no grades given in PBL. This could have led participants to feel less concern for how their behavior would affect their teammates. This is not to say that individuals did not care about their teammates, but it could mean that as an individual studied and prepared for a PBL session, he or she had less concern regarding how his or her actions affected teammates. The participants might have been focused solely on learning the material for themselves and might have given little thought to how they would teach concepts or defend their ideas to their teammates. Conversely, in a PI or a TBL session, participants felt accountable to their classmates and most likely prepared for how they would share information or defend their ideas to their classmates.

My research has provided evidence for the beneficial influence of PBL as an individual learning activity but not as a team learning activity. Individual learning is not a bad thing, but

PBL is structured so that participants will learn from their peers. If participants are not learning from their peers in PBL, there might be more efficacious and cost effective ways for students to learn material individually. However, if grades are added to the structure of PBL, it is reasonable to assume that participants would perceive higher levels of accountability to their teammates, and PBL would then become an effective team learning activity.

Another reason that incorporating grades might be beneficial to PBL relates to team decision making. As previously discussed, including a decision making component to a training intervention has beneficial outcomes. However, the beneficial outcomes associated with team decision making might be greater if feedback is incorporated into the decision making component. According to Hollenbeck, Ilgen, LePine, Colquitt, & Hedlund (1998) group decision making is most effective when it is accompanied by feedback. However, feedback should be presented properly. Earley, Northcraft, Lee, and Lituchy (1990) showed that feedback is very effective when it is presented in terms of both outcomes and processes. In PBL participants receive feedback, particularly from facilitators, concerning processes. But, participants receive little feedback concerning outcomes. Providing feedback in the form of grades would add the outcome element to feedback in PBL. This could add to the benefits associated with team decision making because the participants would receive feedback concerning both outcomes and processes in a PBL setting

Depth of knowledge. The depth to which participants studied a topic influenced participants' understanding of health and biology. Participants used more words pertaining to health when discussing PBL compared to TBL and PI. This was an interesting positive result concerning PBL interventions. Discussion of health in PBL sessions could reflect the single medical issue focus of a PBL session.. For example, a typical PBL case could have revolved around heart attacks or asthma. In contrast, TBL sessions and PI sessions were typically much broader, focusing on topics such as cardiovascular pathophysiology or lung mechanics. Because of the more specific focus in a PBL session, participants were more likely to gain a deeper understanding of the topic. This led to more discussion of health.

Similarly, participants used more words pertaining to biology when discussing PBL compared to TBL and PI. As discussed in the previous paragraph, in PBL sessions, participants went into further depth concerning a topic compared to a TBL session or a PI session. Therefore, they would spend more time and cognitive energy thinking about and discussing biology.

An implication of these findings is that in both PI and TBL sessions the faculty members that conducted the sessions included questions that were written to help participants gain a deeper understanding of the topic. However, the depth of understanding attained in a PI or TBL session does not appear to have equaled the depth of understanding attained in a PBL session.

Relationship between feedback and accountability. For seven of the nine possible tests, feedback and accountability had positive significant relationships with each other. This finding is consistent with prior theoretical reasoning. Though in my study I had no way to determine the direction of this relationship I believe that an increase in feedback lead to an increase in accountability. Similarly, London, Smither and Adsit (1997) claimed that the need to justify a decision increases accountability. In many active learning situations participants feel the need to justify their decisions because they know will receive feedback concerning their decisions in the future. Individuals do not want to provide wrong information or contribute to an incorrect decision because they know they will receive feedback in future that will inform them about the correctness of their decision. Therefore, an increase in feedback should lead to an increase in the desire to be able to justify decisions. This should lead to higher levels of accountability.

Content versus process. Participants focused on different aspects of the active learning interventions during the interviews. When discussing TBL and PI students used more words that related to processes. This is highlighted by the result that participants perceived more words related to insight in TBL and PI than in PBL. In PBL, participants focused more on content, specifically health and biology. Words related to health and biology were words that related to the content of the course and what the students needed to learn in order to succeed.

Distinguishing key components between interventions. Though there were slight mean differences across the three interventions, my results suggested that participants had difficulty distinguishing between different levels of feedback, accountability, or guided exploration across the three active learning interventions. This was a surprising finding due to the very different nature of the structure of each of the three interventions. There are many possible explanations for this phenomenon.

First, participants might not have ever concerned themselves with distinguishing the different interventions. The participants could have just perceived all the interventions as schoolwork or learning activities. However, the data I collected from participant interviews renders this conclusion unlikely.

A second explanation for the lack of distinction could be that participants clumped all these interventions together as “non-lecture” based learning activities. In all of the interviews that I conducted, participants mentioned about how they would like to have lectures returned to their curriculum. Many participants saw lectures as the best source of “primary” learning and these three active learning interventions as “secondary” learning activities.

A third reason is that participants were able to differentiate between the interventions themselves but might have spent little time thinking about what made the interventions different from each other. Based on the interview data, participants were very concerned with how they felt about the efficacy of each intervention but were less concerned about why they felt that one intervention was more efficacious or enjoyable than another. This is understandable due to the nature of the schedule of the participants. All participants were enrolled in medical school, and the amount of information that each participant needed to learn and comprehend in order to excel in medical school was very high. Therefore, participants were not likely to have spent time and cognitive energy considering the differences between interventions due to limited cognitive resources. However, I could have measured constructs better with higher quality survey questions. Else, future researchers could choose a sample that has a higher probability of being able to distinguish between interventions.

Organizational climate. One variable that I did not previously mention or consider as part of this study was the climate of the organization in which this research was conducted. This point is particularly relevant concerning the measurement of the key components and outcomes of PBL. Most participants did not believe PBL added to their learning and desired that the administration do away with PBL altogether. Additionally, four of the six interviewees explicitly stated that they preferred PI over TBL and PBL, and the other two interviewees enjoyed PI and TBL equally. Obviously, this means that no one chose PBL as their preferred method of active learning. This is likely due to the fact that PBL was introduced later in the academic year when the participants were focused on the United States Medical Licensing Examination that would occur a few months after the introduction of PBL. The participants believed that PBL had potential as a learning tool, but the majority of the interviewees expressed a preference to use the time that is spent in PBL sessions to study for the upcoming examination.

This climate could have affected the results in a number of ways. First, when participants were asked about PBL, their negative emotions concerning PBL could have impacted their perceptions of the key components and outcomes. Schwarz (2000) argued that individuals' assessments of their experiences are based on two moments: the moment of peak intensity and the ending. In the case of PBL, both of these moments were likely to have been perceived as negative due to the climate, which could have affected participants' survey responses. Second, participants could have perceived that the surveys that they were given could be used to send a message to the administration. Because the majority of the participants desired that PBL be taken out of the curriculum, they could have answered the survey measures in a way that they thought would influence the administration to adopt their preferences instead of answering honestly.

General Implications

Professionalism is an area of increased interest to both academics and practioners in many fields, including Psychology (Palmero, Janicke, McQuaid, Mullins, Robins, & Wu, 2014) Medicine (Wynia, Papadakis, Sullivan, & Hafferty, 2014), and Business (Bordass & Leaman, 2013). As the world becomes more complex and oversight becomes more difficult, professionalism is an important characteristic that many industries desire in their employees or students. Obviously, the goal of most academics and practioners is to increase the levels of professionalism in their students or trainees. This study provided evidence that including high levels of feedback and accountability might increase levels of professionalism. If academics and practioners desire to increase levels of professionalism, this study provided evidence that suggests that including high levels of feedback and accountability in an intervention might increase levels of professionalism.

Similarly, emotion control is an area of increase interest in Psychology (Teper, Segal, & Inzlicht, 2013), Medicine (Niven, Totterdell, Miles, Webb, & Sheeran, 2013), and Business (Hulsheger, Alberts, Feinholdt, & Lang, 2013). As academic and business environments become increasingly reliant upon teams and group interaction, organizations might have a greater interest in increasing individual's levels of emotion control. Therefore, this study provided evidence that if an organization desires to increase levels of emotion control the organization should include elements of accountability into any intervention that the organization employs to increase emotion control.

I measured professionalism with a shortened version of the scale created by Chisholm, Cobb, Duke, McDuffie, and Kennedy (2006). Their original version consisted of thirty two items whereas the revised version that I used consisted of only six items. The main reason that I revised this measure was to decrease the amount of time that it took to take the measure. Most participants completed the revised measure in less than one minute. Most people do not enjoy taking surveys, and if a shorter version of a measure is available then it would be beneficial to researchers to use the shorter version. However, more validation research is needed before anyone can conclude that the revised version that I used measured professionalism effectively.

Finally, one interesting finding related to change in perceptions and outcomes. Specifically, scores related to feedback, guided exploration, and learning decreased across time. PI was measured during Course 1, PI and TBL were measured during Course 2, and PI, TBL, and PBL were measured during Course 3. Based on these findings, we can conclude that participants perceived that there was more feedback and guided exploration in PI, less feedback and guided exploration in TBL, and even less in PBL. These findings make sense based on the structure of the PI, TBL, and PBL interventions. However, much more research is needed in order to substantiate these claims. Furthermore, it would be unwise to conclude that in PI participants learn better than in TBL and PBL due to the many variables that influence learning that I was unable to account for in this study.

Confounds and Limitations

Confounds. My study had many confounds that make it difficult to reach conclusions. One of these confounds was the fact that the three active learning interventions were not administered separately. PI was the only active learning intervention during Course 1, but participants concurrently participated in both PI and TBL classes during Course 2, and participants concurrently participated in PI, TBL, and PBL classes during Course 3. This occurred due to the nature of the curriculum of the medical school in which I conducted my study. Due to this, it was difficult to conclude that one intervention versus another was the reasons for the observed results for different courses. Though I had theoretical reasoning to justify relationships of PI, TBL, and PBL with outcomes, it was difficult to conclude that these interventions were the reason for the observed results due to the confound between intervention and course.

Another confound that made it difficult to relate components of the active learning conditions to outcomes was that participants were in different courses throughout the year. This lead to several distinct issues. First, in each course, the participants were taught by different instructors. Though some topics that were part of every course (e.g., pharmacology) were always taught by the same instructor, the courses themselves were taught by different instructors. Obviously, this introduces variability in the learning process that I could not control. Different instructors had different amounts of experience and training related to the active learning interventions. Also, each instructor invariably differed in the way he or she did things such as question writing, material delivery, student involvement, and humor. Also, participants tended to prefer some professors more than others, which can influence things such as study time and attendance.

Second, because participants were in different courses, participants learned different material during each course. Possibly, some material was more difficult to learn than other material for the class as a whole and for individual participants. This was especially problematic as I assessed learning as an outcome. I used final exam scores as the measure of learning. I observed an eleven point decrease between Course 1 and Course 2 in exam scores. Then, I observed a ten point increase between Course 2 and Course 3 in exam scores. These differences might be due to the nature of the material that was tested on the exams, but I do not have a definitive way to know that.

Third, though active learning techniques were the main method of content delivery for each course, participants had additional learning opportunities, which could have affected outcomes. Participants were exposed activities such as lectures, online videos, and study groups. All of these activities could have affected outcomes but were not accounted for in this study.

Fourth, the courses varied in length. That is, not all of the courses covered the same time period. For example, Course 2 was only two weeks long whereas Course 1 was five weeks long. This confound introduced variability into the amount of time each participant spent in the active learning conditions. This could have led to differential effects on outcomes.

A final confound was that different professors structured their courses differently. This affected the amount of active learning interventions that participants participated in. For example, the professor for Course 1 structured the course so that participants were involved in a PI session almost every day of the work week. In contrast, during Course 2, participants were involved in active learning interventions about two times per week. By participating in active learning interventions for different amounts of time, participants might not have had the requisite time to improve outcomes from one course compared to another course.

Limitations. This study had limitations due to situational constraints that could not be changed or controlled for. First, the sample size was limited. If every student possible participated in this study, the total sample size would have been one hundred and nine participants. According to many researchers (e.g., Hsieh, Bloch, & Larsen, 1998), this sample size is not adequate to conduct linear regression. Due to the limited sample size, it would not be advisable to generalize the results of this study outside of a medical school setting. Research on key components of active learning needs to be conducted with larger samples in different settings to determine what components of active learning account for the positive results. As more and more universities and organizations adopt active learning techniques, research samples should become more readily available to interested researchers.

A further limitation is that of range restriction. The participants in this study were all high achieving medical school students who most likely were already high in self-efficacy, metacognition, emotion control, and professionalism. Range restriction causes a number of statistical issues including a reduction in observed correlations (Sackett & Yang, 2000). This is problematic in this study because the tests of my main hypotheses were bivariate correlations. Furthermore, it was difficult to assess learning across time due to changes in course material and changes in instructor. Though most instructors were trained on how to conduct these three interventions individual differences still occurred.

Conclusion

The purpose of this study was to examine whether feedback, accountability, and guided exploration were key components of active learning. My results did not consistently support my hypotheses. However, I did observe a pattern of differences concerning feedback and accountability in the predicted direction in all three interventions. Also, feedback had a positive relationship with professionalism in both PI and PBL, and accountability had positive relationships with emotion control and professionalism in both PI and TBL. Furthermore, my results highlighted issues related to each key component. First, results concerning perceptions of feedback were influenced by the nature of the questions that were chosen, interactions between participants and peers, and the source of feedback. Second, results concerning perceptions of accountability were influenced by team membership. Third, results concerning guided exploration highlighted the importance of adequate measurement. Overall, this study adds to the scientific literature by raising questions regarding which components of active learning have relationships with important outcomes, and what issues affect key components of active learning.

References

Azevedo, R., & Bernard, R. M. (1995). A meta-analysis of the effects of feedback in computer-based instruction. *Journal of Educational Computing Research*, *13*(2), 111-127.

Bakker, A. B., Demerouti, E., & Lieke, L. (2012). Work engagement, performance, and active learning: The role of conscientiousness. *Journal of Vocational Behavior*, *80*(2), 555-564.

Baldwin, T. T., & Magjuka, R. J. (1997). Training as an organizational episode: Pretraining influences on trainee motivation. *Improving training effectiveness in work organizations*, 99-127.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological* review, 84(2), 191.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. Macmillan.

Baumeister, R. F. (1982). A self-presentational view of social phenomena.*Psychological bulletin*, *91*(1), 3.

Beatty, S. J., Kelley, K. A., Metzger, A. H., Bellebaum, K. L., & McAuley, J. W. (2009). Team-based learning in therapeutics workshop sessions. *American journal of pharmaceutical education*, *73*(6).

Bell, B. S., & Kozlowski, S. W. (2008). Active learning: effects of core training design elements on self-regulatory processes, learning, and adaptability. *Journal of Applied psychology*, *93*(2), 296.

Bell, B. S., & Kozlowski, S. W. (2010). Toward a theory of learner-centered training design: An integrative framework of active learning. *Learning, training, and development in organizations*, 263-300.

Benware, C. A., & Deci, E. L. (1984). Quality of learning with an active versus passive motivational set. *American Educational Research Journal*, *21*(4), 755-765.

Bliese, P. D., & Ployhart, R. E. (2002). Growth modeling using random coefficient models: Model building, testing, and illustrations. *Organizational Research Methods*, *5*(4), 362-387.

Bonwell, C. C., & Eison, J. A. (1991). *Active Learning: Creating Excitement in the Classroom. 1991 ASHE-ERIC Higher Education Reports*. ERIC Clearinghouse on Higher Education, The George Washington University, One Dupont Circle, Suite 630, Washington, DC 20036-1183.

Bordass, B., & Leaman, A. (2013). A new professionalism: remedy or fantasy?. *Building Research & Information*, *41*(1), 1-7.

Brady, M., Seli, H., & Rosenthal, J. (2013). “Clickers” and metacognition: A quasi-experimental comparative study about metacognitive self-regulation and use of electronic feedback devices. *Computers & Education*, *65*, 56-63.

Brinkman, W. B., Geraghty, S. R., Lanphear, B. P., Khoury, J. C., del Rey, J. A. G., DeWitt, T. G., & Britto, M. T. (2007). Effect of multisource feedback on resident communication skills and professionalism: a randomized controlled trial. *Archives of pediatrics & adolescent medicine*, *161*(1), 44-49.

Brown, A. L. (1975). The development of memory: Knowing, knowing about knowing, and knowing how to know. *Advances in child development and behavior*, *10(10)*, 103-152.

Bruner, J. S. (1961). The act of discovery. *Harvard educational review*.

Button, S. B., Mathieu, J. E., & Zajac, D. M. (1996). Goal orientation in organizational research: A conceptual and empirical foundation. *Organizational behavior and human decision processes, 67(1),* 26-48.

Burke, L. A., & Hutchins, H. M. (2008). A study of best practices in training transfer and proposed model of transfer. *Human resource development quarterly, 19(2),* 107.

Cohn, D., Atlas, L., & Ladner, R. (1994). Improving generalization with active learning. *Machine learning, 15(2),* 201-221.

Cook, D. A., Levinson, A. J., & Garside, S. (2010). Time and learning efficiency in internet-based learning: A systematic review and meta-analysis. *Advances In Health Sciences Education*, *15(5),* 755-770.

Chiaburu, D. S., & Lindsay, D. R. (2008). Can do or will do? The importance of self-efficacy and instrumentality for training transfer. *Human Resource Development International*, *11(2),* 199-206.

Chisholm, M. A., Cobb, H., Duke, L., McDuffie, C., & Kennedy, W. K. (2006). Development of an instrument to measure professionalism. *American journal of pharmaceutical education*, *70(4).*

Chillarege, K. A., Nordstrom, C. R., & Williams, K. B. (2003). Learning from our mistakes: Error management training for mature learners. *Journal of business and psychology*, *17(3),* 369-385.

Chung, E. K., Rhee, J. A., & Baik, Y. H. (2009). The effect of team-based learning in medical ethics education. *Medical Teacher, 31*(11), 1013-1017.

Craig, R. C. (1953). The transfer value of guided learning. Oxford, England: Bureau of Publications, Teachers Co.

Craig, R. C. (1956). Directed versus independent discovery of established relations. *Journal Of Educational Psychology*, *47(4)*, 223-234.

Crouch, C. H., & Mazur, E. (2001). Peer instruction: Ten years of experience and results. *American journal of physics*, *69(9)*, 970-977.

Crouch, C. H., Watkins, J., Fagen, A. P., & Mazur, E. (2007). Peer instruction: Engaging students one-on-one, all at once. *Research-Based Reform of University Physics*, 1(1), 40-95.

Cunningham, D., & Duffy, T. (1996). Constructivism: Implications for the design and delivery of instruction. *Handbook of research for educational communications and technology*, 170-198.

Dalgarno, B., Kennedy, G., & Bennett, S. (2014). The impact of students' exploration strategies on discovery learning using computer-based simulations. *Educational Media International*, *51*(4), 310-329.

Debowski, S., Wood, R. E., & Bandura, A. (2001). Impact of guided exploration and enactive exploration on self-regulatory mechanisms and information acquisition through electronic search. *Journal of Applied Psychology*, *86(6)*, 1129.

Deci, E. L. (1971). Effects of externally mediated rewards on intrinsic motivation. *Journal of personality and Social Psychology, 18(1),* 105.

Dewhurst, S. A., & Anderson, S. J. (1999). Effects of exact and category repetition in true and false recognition memory. *Memory & Cognition*, *27*(4), 665-673.

Dolmans, D., Michaelsen, L., Van Merrienboer, J., & van der Vleuten, C. (2015). Should we choose between problem-based learning and team-based learning? No, combine the best of both worlds!. *Medical teacher*, 3*7(4)*, 354-359.

Dufresne, R. J., Gerace, W. J., Leonard, W. J., Mestre, J. P., & Wenk, L. (1996). Classtalk: A classroom communication system for active learning. *Journal of computing in higher education*, *7(2)*, 3-47.

Dusterhoff, C., Cunningham, J. B., & MacGregor, J. N. (2014). The effects of performance rating, leader–member exchange, perceived utility, and organizational justice on performance appraisal satisfaction: Applying a moral judgment perspective. *Journal Of Business Ethics*, *119(2)*, 265-273.

Dweck, C. S. (1975). The role of expectations and attributions in the alleviation of learned helplessness. *Journal of personality and social psychology*, *31(4)*, 674.

Dweck, C. S. (1986). Motivational processes affecting learning. *American psychologist*, *41(10)*, 1040.

Earley, P. C., Northcraft, G. B., Lee, C., & Lituchy, T. R. (1990). Impact of process and outcome feedback on the relation of goal setting to task performance. *Academy of Management Journal*, *33*(1), 87-105.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American psychologist*, *34(10)*, 906.

Ford, J. K., Smith, E. M., Weissbein, D. A., Gully, S. M., & Salas, E. (1998). Relationships of goal orientation, metacognitive activity, and practice strategies with learning outcomes and transfer. *Journal of applied psychology*, *83*(2), 218.

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.

Freeman, S., O'Connor, E., Parks, J. W., Cunningham, M., Hurley, D., Haak, D., Dirks, D., & Wenderoth, M. P. (2007). Prescribed active learning increases performance in introductory biology. *CBE-Life Sciences Education*, *6*(2), 132-139.

Frese, M., Brodbeck, F., Heinbokel, T., Mooser, C., Schleiffenbaum, E., & Thiemann, P. (1991). Errors in training computer skills: On the positive function of errors. *Human-Computer Interaction*, *6*(1), 77-93.

Giessner, S. R., van Knippenberg, D., van Ginkel, W., & Sleebos, E. (2013). Team-oriented leadership: The interactive effects of leader group prototypicality, accountability, and team identification. *Journal of Applied Psychology*, *98*(4), 658.

Gist, M. E., & Mitchell, T. R. (1992). Self-efficacy: A theoretical analysis of its determinants and malleability. *The Academy Of Management Review*, *17*(2), 183-211.

Goldstein, I. L., & Ford, J. K. (2002). *Training in Organizations* (4th ed.). Belmont, CA: Wadsworth.

Gray, J. A. (1991). Fear, panic, and anxiety: What's in a name?. *Psychological Inquiry*, *2*(1), 77-78.

Greller, M. M. (1978). The nature of subordinate participation in the appraisal interview. *Academy of Management*

Griffin, M. M., & Griffin, B. W. (1998). An investigation of the effects of reciprocal peer tutoring on achievement, self-efficacy, and test anxiety. *Contemporary Educational Psychology*, *23*(3), 298-311.

Gully, S. M., Payne, S. C., Koles, K., & Whiteman, J. A. K. (2002). The impact of error training and individual differences on training outcomes: an attribute-treatment interaction perspective. *Journal of Applied Psychology*, *87*(1), 143.

Haidet, P., O'Malley, K. J., & Richards, B. (2002). An initial experience with “team learning” in medical education. Academic Medicine, 77(1), 40-44.

Harrison, D. A., Mohammed, S., McGrath, J. E., Florey, A. T., & Vanderstoep, S. W. (2003). Time matters in team performance: Effects of member familiarity, entrainment, and task discontinuity on speed and quality.*Personnel Psychology*, *56*(3), 633-669.

Hatano, G., & Inagaki, K. (1986). Two courses of expertise. RESEARCH AND CLINICAL CENTER FOR CHILD DEVELOPMENT Annual Report, 6: 27-36

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of educational research*, *77*(1), 81-112.

Henderson, E. N. (1903). A study of memory: For connected trains of thought. The Psychological Review: Monograph Supplements, 5(6).

Heinz Jr, P., & Steele-Johnson, D. (2004). Clarifying the conceptual definitions of goal orientation dimensions: Competence, control, and evaluation. *Organizational Analysis*, *12*(1), 5-19.

Hinkin, T. R. (1995). A review of scale development practices in the study of organizations. *Journal of management*, *21*(5), 967-988.

Hochwarter, W. A., Perrewé, P. L., Hall, A. T., & Ferris, G. R. (2005). Negative affectivity as a moderator of the form and magnitude of the relationship between felt accountability and job tension. *Journal Of Organizational Behavior, 26*(5), 517-534.

Hogg, M. A., & Williams, K. D. (2000). From I to we: Social identity and the collective self. *Group Dynamics: Theory, Research, And Practice, 4*(1), 81-97.

Hollenbeck, J. R., Ilgen, D. R., LePine, J. A., Colquitt, J. A., & Hedlund, J. (1998). Extending the multilevel theory of team decision making: Effects of feedback and experience in hierarchical teams. *Academy of Management Journal*, *41*(3), 269-282.

Hopkins, K. D., Oldridge, O. A., & Williamson, M. L. (1965). An empirical comparison of pupil achievement and other variables in graded and ungraded classes. *American Educational Research Journal*, *2*(4), 207-215.

Hounsell, D. (2007). Towards more sustainable feedback to students.*Rethinking assessment in higher education*, 101-113.

Hsieh, F. Y., Bloch, D. A., & Larsen, M. D. (1998). A simple method of sample size calculation for linear and logistic regression. *Statistics in medicine*, *17*(14), 1623-1634.

Huffaker, D. A., & Calvert, S. L. (2003). The new science of learning: Active learning, metacognition, and transfer of knowledge in e-learning applications. *Journal of Educational Computing Research*, *29*(3), 325-334.

Hung, W., Jonassen, D. H., & Liu, R. (2008). Problem-based learning. *Handbook of research on educational communications and technology*, 3, 485-506.

Hülsheger, U. R., Alberts, H. J., Feinholdt, A., & Lang, J. W. (2013). Benefits of mindfulness at work: the role of mindfulness in emotion regulation, emotional exhaustion, and job satisfaction. *Journal of Applied Psychology*,*98*(2), 310.

Husen, T. (1968). Lifelong learning in the ‘Educative Society’. *Applied Psychology*, 17(2), 87-98.

Ilgen, D. R., Fisher, C. D., & Taylor, M. S. (1979). Consequences of individual feedback on behavior in organizations. *Journal of applied psychology*, *64*(4), 349.

Iran-Nejad, A. (1990). Active and dynamic self-regulation of learning processes. *Review Of Educational Research*, 60(4), 573-602.

Jahromi, M. K., & Mosallanejad, L. (2014). The impact of reality therapy on metacognition, stress and hope in addicts. *Global journal of health science*,*6*(6), 281.

Kanfer, R., & Ackerman, P. L. (1990). Ability and metacognitive determinants of skill acquisition and transfer. MINNESOTA UNIV MINNEAPOLIS DEPT OF PSYCHOLOGY.

Keeler, C. M., & Steinhorst, R. K. (1995). Using small groups to promote active learning in the introductory statistics course: A report from the field. *Journal of Statistics Education*, 3(2), 1-8.

Keith, N. (2005). *Self-regulatory processes in error management training* (Doctoral dissertation). Retrieved from Universitätsbibliothek Giessen.

Keith, N., & Frese, M. (2005). Self-regulation in error management training: emotion control and metacognition as mediators of performance effects.*Journal of Applied Psychology*, *90*(4), 677.

Keith, N., & Frese, M. (2008). Effectiveness of error management training: a meta-analysis. *Journal of Applied Psychology*, *93*(1), 59.

Kittell, J. E. (1957). An experimental study of the effect of external direction during learning on transfer and retention of principles. *Journal of Educational Psychology*, *48*(7), 391.

Kim, J. H., & Lee, C. (2001). Implications of near and far transfer of training on structured on-the-job training. *Advances in Developing Human Resources*, *3*(4), 442-451.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational psychologist, 41*(2), 75-86.

Klašnja-Milićević, A., Vesin, B., Ivanović, M., & Budimac, Z. (2011). E-Learning personalization based on hybrid recommendation strategy and learning style identification. *Computers & Education, 56*(3), 885-899.

Klein, D. F. (2002). Historical aspects of anxiety. *Dialogues in Clinical Neuroscience, 4*(3), 295–304.

Kluge, A., Sauer, J., Burkolter, D., & Ritzmann, S. (2010). Designing training for temporal and adaptive transfer: A comparative evaluation of three training methods for process control tasks. *Journal Of Educational Computing Research*, *43*(3), 327-353.

Koles, P. G., Stolfi, A., Borges, N. J., Nelson, S., & Parmelee, D. X. (2010). The impact of team-based learning on medical students' academic performance. *Academic Medicine, 85*(11), 1739-1745.

Koohang, A., & Paliszkiewicz, J. (2013). Knowledge construction in e-learning: An empirical validation of an active learning model. *The Journal of Computer Information Systems, 53*(3), 109.

Kong, L. N., Qin, B., Zhou, Y. Q., Mou, S. Y., & Gao, H. M. (2014). The effectiveness of problem-based learning on development of nursing students' critical thinking: A systematic review and meta-analysis. *International journal of nursing studies, 51*(3), 458-469.

Kouchaki, M., & Desai, S. D. (2014). Anxious, threatened, and also unethical: How anxiety makes individuals feel threatened and commit unethical acts.

Kozlowski, S. W., & Bell, B. S. (2006). Disentangling achievement orientation and goal setting: effects on self-regulatory processes. *Journal of Applied Psychology*, *91*(4), 900.

Kozlowski, S. W., Gully, S. M., Brown, K. G., Salas, E., Smith, E. M., & Nason, E. R. (2001). Effects of training goals and goal orientation traits on multidimensional training outcomes and performance adaptability. *Organizational behavior and human decision processes*, *85*(1), 1-31.

Kozlowski, S. W., Toney, R. J., Mullins, M. E., Weissbein, D. A., Brown, K. G., & Bell, B. S. (2001). Developing adaptability: A theory for the design of integrated-embedded training systems. *Advances in human performance and cognitive engineering research*, *1*, 59-124.

Kraiger, K., & Jung, K. (1997). Linking Objectives to Evaluation Criteria. In Training for a Rapidly Changing Workplace (p. 152). Washington D.C.: American Psychological Association.

Kraiger, K., Ford, J. K., & Salas, E. (1993). Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation. *Journal of applied psychology*, *78*(2), 311.

Lerner, J. S., & Tetlock, P. E. (1999). Accounting for the effects of accountability. *Psychological bulletin*, *125*(2), 255.

Lewin, K. (1947). Frontiers in group dynamics II. Channels of group life; social planning and action research. *Human relations*, *1*(2), 143-153.

Lilly, R. C. (1951). An investigation of problems in adult education, with special emphasis upon a particular program in the state of Virginia.

Linton, D. L., Pangle, W. M., Wyatt, K. H., Powell, K. N., & Sherwood, R. E. (2014). Identifying Key Features of Effective Active Learning: The Effects of Writing and Peer Discussion. *CBE-Life Sciences Education*, *13*(3), 469-477.

London, M., Smither, J. W., & Adsit, D. J. (1997). Accountability The Achilles' Heel of Multisource Feedback. *Group & Organization Management*,*22*(2), 162-184.

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*(5), 368-376.

Martin, T., Rivale, S. D., & Diller, K. R. (2007). Comparison of student learning in challenge-based and traditional instruction in biomedical engineering. *Annals of biomedical engineering*, *35*(8), 1312-1323.

Massenberg, A. C., Spurk, D., & Kauffeld, S. (2015). Social support at the workplace, motivation to transfer and training transfer: a multilevel indirect effects model. *International Journal of Training and Development.*

Mathieu, J. E., & Martineau, J. W. (1997). Individual and situational influences on training motivation. *Improving training effectiveness in work organizations*,*193*, 221.

Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning?. *American psychologist*, *59*(1), 14.

Mayer, R. E., & Greeno, J. G. (1972). Structural differences between outcomes produced by different instructional methods. *Journal of educational psychology*, 63(2), 165.

McCarthy, J. P., & Anderson, L. (2000). Active learning techniques versus traditional teaching styles: two experiments from history and political science. *Innovative Higher Education*, *24*(4), 279-294.

McConnell, J. J. (1996). Active learning and its use in computer science. *ACM SIGCSE Bulletin*, *28*(SI), 52-54.

McDaniel, M. A., & Schlager, M. S. (1990). Discovery learning and transfer of problem-solving skills. *Cognition and Instruction*, *7*(2), 129-159.

McParland, M., Noble, L. M., & Livingston, G. (2004). The effectiveness of problem-based learning compared to traditional teaching in undergraduate psychiatry. *Medical education*, *38*(8), 859-867.

Mead, S., Hilton, D., & Curtis, L. (2001). Peer support: a theoretical perspective. *Psychiatric rehabilitation journal*, *25*(2), 134.

Melka, S. E., Lancaster, S. L., Bryant, A. R., & Rodriguez, B. F. (2011). Confirmatory factor and measurement invariance analyses of the emotion regulation questionnaire. *Journal of clinical psychology*, *67*(12), 1283-1293.

Mezulis, A. H., Abramson, L. Y., Hyde, J. S., & Hankin, B. L. (2004). Is there a universal positivity bias in attributions? A meta-analytic review of individual, developmental, and cultural differences in the self-serving attributional bias. *Psychological bulletin*, *130*(5), 711.

Michaelsen, L. K., Watson, W., Cragin, J. P., & Fink, L. D. (1982). Team learning: A potential solution to the problems of large classes. Journal of Management Education, 7(1), 13-22.

Michaelsen, L. K., Sweet, M., & Parmelee, D. X. (Eds.). (2008). *Team-Based Learning: Small Group Learning's Next Big Step: New Directions for Teaching and Learning, Number 116* (Vol. 103). John Wiley & Sons.

Miflin, Campbell, Price, & Miflin, B. (2000). A conceptual framework to guide the development of self-directed, lifelong learning in problem-based medical curricula. *Medical Education,34*(4), 299-306.

Meyers, C., & Jones, T. B. (1993). Promoting Active Learning. Strategies for the College Classroom. Jossey-Bass Inc., Publishers, 350 Sansome Street, San Francisco, CA 94104.

Naveh, E., Katz-Navon, T., & Stern, Z. (2015). Active learning climate and employee errors: The moderating effects of personality traits. *Journal of Organizational Behavior*, *36*(3), 441-459.

Neufeld, V. R., & Barrows, H. S. (1974). The 'McMaster Philosophy': An approach to medical education. Journal Of Medical Education, 49(11), 1040-1050.

Neville, A. J. (2009). Problem-based learning and medical education forty years on. Medical Principles and Practice, 18(1), 1-9.

Newmann, F., King, M. B., & Rigdon, M. (1997). Accountability and school performance: Implications from restructuring schools. *Harvard educational review*, *67*(1), 41-75.

Nicholls, J. G. (1975). Causal attributions and other achievement-related cognitions: Effects of task outcome, attainment value, and sex. *Journal of Personality and Social Psychology*, *31*(3), 379.

Niven, K., Totterdell, P., Miles, E., Webb, T. L., & Sheeran, P. (2013). Achieving the same for less: Improving mood depletes blood glucose for people with poor (but not good) emotion control. *Cognition & emotion*, *27*(1), 133-140.

Nordstrom, C. R., Wendland, D., & Williams, K. B. (1998). “To err is human”: An examination of the effectiveness of error management training. *Journal of business and psychology*, *12*(3), 269-282.

Olson, S., & Riordan, D. G. (2012). Engage to Excel: Producing One Million Additional College Graduates with Degrees in Science, Technology, Engineering, and Mathematics. Report to the President. Executive Office of the President.

Palermo, T. M., Janicke, D. M., McQuaid, E. L., Mullins, L. L., Robins, P. M., & Wu, Y. P. (2014). Recommendations for training in pediatric psychology: Defining core competencies across training levels. *Journal of pediatric psychology*, *39*(9), 965-984.

Palincsar, A. S., & Brown, D. A. (1987). Enhancing instructional time through attention to metacognition. *Journal Of Learning Disabilities, 20*(2), 66-75.

Paris, S. G., & Winograd, P. (1990). How metacognition can promote academic learning and instruction. *Dimensions of thinking and cognitive instruction*, *1*, 15-51.

Payne, S. C., Youngcourt, S. S., & Beaubien, J. M. (2007). A meta-analytic examination of the goal orientation nomological net. *Journal of Applied Psychology*, *92*(1), 128.

Phil, C. (2000). Reaffirming a proud tradition: Universities and lifelong learning. *Active Learning In Higher Education, 1*(2), 101-125.

Prince, M. (2004). Does active learning work? A review of the research. JOURNAL OF ENGINEERING EDUCATION-WASHINGTON-, 93, 223-232.

Quiñones, M. A. (1997). Contextual influences on training effectiveness.*Training for a rapidly changing workplace: Applications of psychological research*, 177-199.

Rawsthorne, L. J., & Elliot, A. J. (1999). Achievement goals and intrinsic motivation: A meta-analytic review. *Personality and Social Psychology Review*,*3*(4), 326-344.

Riggs, M. L., Warka, J., Babasa, B., Betancourt, R., & Hooker, S. (1994). Development and validation of self-efficacy and outcome expectancy scales for job-related applications. *Educational and psychological measurement, 54(3),* 793-802.

Rohn, D., Austin, J., & Lutrey, S. M. (2003). Using feedback and performance accountability to decrease cash register shortages. *Journal of Organizational Behavior Management*, *22*(1), 33-46.

Rosenzweig, S. (1945). The picture-association method and its application in a study of reactions to frustration. *Journal of personality*, *14*(1), 3-23.

Rothman, D. J. (1991). Strangers at the bedside: a history of how law and bioethics transformed medical decision making. Basic Books.

Royle, M. T., Hall, A. T., Hochwarter, W. A., Perrewé, P. L., & Ferris, G. R. (2005). THE INTERACTIVE EFFECTS OF ACCOUNTABILITY AND JOB SELF-EFFICACY ON ORGANIZATIONAL CITIZENSHIP BEHAVIOR AND POLITICAL BEHAVIOR. *Organizational Analysis*, *13*(1).

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist*, *55*(1), 68.

Sackett, P. R., & Yang, H. (2000). Correction for range restriction: an expanded typology. *Journal of Applied Psychology*, *85*(1), 112.

Salasoo, A., Shiffrin, R. M., & Feustel, T. C. (1985). Building permanent memory codes: codification and repetition effects in word identification.*Journal of Experimental Psychology: General*, *114*(1), 50.

Saito, P. T., de Rezende, P. J., Falcao, A. X., Suzuki, C. T., & Gomes, J. F. (2014). An active learning paradigm based on a priori data reduction and organization. *Expert Systems with Applications*, *41*(14), 6086-6097.

Sampson, D. C., & Jackson, M. J. (2014). Armchair Travel: An Active Learning Approach to Increasing Global Awareness and Participant Self-Efficacy.*Developments in Business Simulation and Experiential Learning*, *34*.

Sanson-Fisher, R. W., & Lynagh, M. C. (2005). Problem-based learning: a dissemination success story?. *Medical Journal of Australia*, *183*(5), 258.Savery, J. R. (2015). Overview of problem-based learning: Definitions and distinctions. *Essential Readings in Problem-Based Learning: Exploring and Extending the Legacy of Howard S. Barrows*, 5.

Savery, J. R. (2015). Overview of problem-based learning: Definitions and distinctions. *Essential readings in problem-based learning: Exploring and extending the legacy of Howard S. Barrows*, 5-15.

Schermerhorn, S. M., Goldschmid, M. L., & Shore, B. M. (1975). Learning basic principles of probability in student dyads: A cross-age comparison. *Journal of Educational Psychology*, *67*(4), 551.

Schunk, D. H. (1983). Ability versus effort attributional feedback: Differential effects on self-efficacy and achievement. *Journal of educational psychology*,*75*(6), 848.

Schwarz, N. (2000). Emotion, cognition, and decision making. *Cognition & Emotion*, *14*(4), 433-440.

Scott, M. B., & Lyman, S. M. (1968). Accounts. *American sociological review*, 46-62.

Shin, J. H., Haynes, R. B., & Johnston, M. E. (1993). Effect of problem-based, self-directed undergraduate education on life-long learning. *CMAJ: Canadian Medical Association Journal*, *148*(6), 969.

Shin, I., & Kim, J. (2013). The effect of problem-based learning in nursing education: A meta-analysis. *Advances In Health Sciences Education*, 18(5), 1103-1120.

Silberman, M. (1996). Active Learning: 101 Strategies To Teach Any Subject. Prentice-Hall, PO Box 11071, Des Moines, IA 50336-1071.

Sisk, R. J. (2011). Team-based learning: systematic research review. *The Journal of nursing education, 50*(12), 665-669.

Smith, T. L. (1896). On muscular memory. *The American Journal of Psychology*, *7*(4), 453-490.

Steele-Johnson, D., & Kalinoski, Z. T. (2014). Error Framing Effects on Performance: Cognitive, Motivational, and Affective Pathways. *The Journal of psychology*, *148*(1), 93-111.

Steelman, L. A., Levy, P. E., & Snell, A. F. (2004). The feedback environment scale: Construct definition, measurement, and validation. *Educational and Psychological Measurement, 64(1)*, 165-184.

Stefanou, C., Stolk, J. D., Prince, M., Chen, J. C., & Lord, S. M. (2013). Self-regulation and autonomy in problem-and project-based learning environments. *Active Learning in Higher Education*, *14*(2), 109-122.

Sitzmann, T., & Ely, K. (2011). A meta-analysis of self-regulated learning in work-related training and educational attainment: what we know and where we need to go. *Psychological bulletin*, *137*(3), 421.

Slavin, R. E. (2011). Cooperative learning. *Learning and Cognition in Education Elsevier Academic Press, Boston*, 160-166.

Sweller, J., Mawer, R. F., & Ward, M. R. (1983). Development of expertise in mathematical problem solving. *Journal of Experimental Psychology: General*, *112*(4), 639.

Sung, H. Y., Hwang, G. J., & Yen, Y. F. (2015). Development of a contextual decision-making game for improving students' learning performance in a health education course. *Computers & Education, 82*, 179-190.

Taris, T. W., Kompier, M. A., De Lange, A. H., Schaufeli, W. B., & Schreurs, P. J. (2003). Learning new behaviour patterns: a longitudinal test of Karasek's active learning hypothesis among Dutch teachers. *Work & Stress*, *17*(1), 1-20.

Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of language and social psychology*, *29*(1), 24-54.

Teper, R., Segal, Z. V., & Inzlicht, M. (2013). Inside the mindful mind how mindfulness enhances emotion regulation through improvements in executive control. *Current Directions in Psychological Science*, *22*(6), 449-454.

Thomas, P., Dose, J. J., & Scott, K. S. (2002). Relationships between accountability, job satisfaction, and trust. *Human Resource Development Quarterly*, 13(3), 307-323.

Thompson, C. A., Califf, M. E., & Mooney, R. J. (1999, June). Active learning for natural language parsing and information extraction. In ICML (pp. 406-414).

Thorndike, E. L. (1908). Memory for paired associates. *Psychological Review*, *15*(2), 122.

Tuckman, B. W. (1965). Developmental sequence in small groups. *Psychological bulletin*, *63*(6), 384.

Tuovinen, J. E., & Sweller, J. (1999). A comparison of cognitive load associated with discovery learning and worked examples. *Journal of educational psychology*, *91*(2), 334.

Van den Bergh, L., Ros, A., & Beijaard, D. (2013). Teacher feedback during active learning: Current practices in primary schools. *British Journal of Educational Psychology*, *83*(2), 341-362.

VandeWalle, D. (1997). Development and validation of a work domain goal orientation instrument. *Educational and Psychological Measurement*, *57*(6), 995-1015.

Vasan, N. S., DeFouw, D. O., & Compton, S. (2009). A survey of student perceptions of team-based learning in anatomy curriculum: Favorable views unrelated to grades. *Anatomical sciences education, 2*(4), 150.

Vasan, N. S., DeFouw, D. O., & Holland, B. K. (2008). Modified use of team-based learning for effective delivery of medical gross anatomy and embryology. *Anatomical sciences education, 1*(1), 3-9.

Von Glasersfeld, E. (1989). Cognition, construction of knowledge, and teaching. *Syntheses, 80*(1), 121-140

Waldman, D. A., Atwater, L. E., & Antonioni, D. (1998). Has 360 degree feedback gone amok?. *The Academy of Management Executive*, *12*(2), 86-94.

Wells, A., & Cartwright-Hatton, S. (2004). A short form of the metacognitions questionnaire: Properties of the MCQ-30. *Behaviour Research And Therapy, 42(4),* 385-396.

Westman, M., Eden, D., & Shirom, A. (1985). Job stress, cigarette smoking and cessation: the conditioning effects of peer support. *Social science & medicine, 20*(6), 637-644.

White, R. W. (1959). Motivation reconsidered: the concept of competence. *Psychological review*, *66*(5), 297.

Woike, B. A. (1995). Most-memorable experiences: Evidence for a link between implicit and explicit motives and social cognitive processes in everyday life. *Journal of Personality and Social Psychology*, *68*(6), 1081.

Wolfe, L. M. (1963). Lifelong learning and adjustment in the later years. *Adult Education Quarterly, 14*(1), 26-32.

Wimmers, P. F., & Lee, M. (2015). Identifying longitudinal growth trajectories of learning domains in problem-based learning: a latent growth curve modeling approach using SEM. *Advances in Health Sciences Education*, *20*(2), 467-478.

Wynia, M. K., Papadakis, M. A., Sullivan, W. M., & Hafferty, F. W. (2014). More than a list of values and desired behaviors: A foundational understanding of medical professionalism. *Academic Medicine*, *89*(5), 712-714.

Yoder, J. D., & Hochevar, C. M. (2005). Encouraging active learning can improve students' performance on examinations. *Teaching of Psychology*, *32*(2), 91-95.

Young, M. R. (2005). The motivational effects of the classroom environment in facilitating self-regulated learning. *Journal of Marketing Education*, *27*(1), 25-40.

Zingaro, D. (2014, March). Peer instruction contributes to self-efficacy in CS1. In *Proceedings of the 45th ACM technical symposium on Computer Science Education* (pp. 373-378). ACM.

Appendix A

Accountability

INSTRUCTIONS: Below are statements describing people's classroom behaviors. Please use the following scale to indicate how accurate each statement reflects **your own** behavioral tendencies.

1	2	3	4	5
Not at All		Neutral		To a Great Extent

1. To what extent are your classmates interested in how well you perform your schoolwork?
2. To what extent are your classmates interested in the method you use to perform your classwork?
3. To what extent does your level of performance of your classwork have an impact on your classmates?
4. To what extent are your classmates aware of the methods you use to perform your classwork?
5. To what extent are your classmates aware of the effectiveness of your performance in your classwork?
6. To what extent do you have to justify the methods that you use in performing your classwork to your classmates?
7. To what extent do the methods that you use to perform your classwork have an impact on your classmates?
8. To what extent do you have to justify your effectiveness in performing your classwork to your coworkers?
9. In performing your classwork, to what extent are you consciously aware of the concerns of your classmates?

Adapted from :

Thomas, P., Dose, J. J., & Scott, K. S. (2002). Relationships between accountability, job satisfaction, and trust. *Human Resource Development Quarterly*, 13(3), 307-323.

Appendix B

Feedback Quality

INSTRUCTIONS: Below are statements describing feedback. Please use the following scale to indicate how accurate each statement reflects **your own** perceptions about the feedback you received.

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

1. My classmates (instructor) give me useful feedback about my classroom performance.
2. The performance feedback I receive from my classmates (instructor) is helpful.
3. I value the feedback I receive from my classmates (instructor).
4. The feedback I received from my classmates (instructor) helps me do my classwork.
5. The performance information I receive from my classmates (instructor) is generally not very meaningful. (Reversed)

Adapted from:

Steelman, L. A., Levy, P. E., & Snell, A. F. (2004). The feedback environment scale: Construct definition, measurement, and validation. *Educational and Psychological Measurement, 64(1)*, 165-184.

Appendix C

Guided Exploration

1	2	3	4	5	6	7
Not at all			Neutral			To a great extent

1. Rate the extent to which a student’s exploration is guided in the ___ course.
2. To what extent is a student able to explore information related to this course without any suggestions or guidance?
3. To what extent does the course instructor offer suggestions to students regarding the process they should use to obtain information relevant to TBL, PBL, or peer instruction assignments in this course?
4. To what extent does the course instructor offer suggestions to students regarding the type of content students should look for outside of class in order to complete TBL, PBL, or peer instruction assignments?
5. To what extent do you have flexibility with what you can study outside of the classroom?
6. To what extent are you given specific instructions about what to study in this course?

Appendix D

Demographics

Age: ________(State your age in years)

Gender (Circle one): Male Female

Ethnicity:	African American	Asian	Hispanic	Native American
(Circle one)	Pacific Islander	White/Caucasian		Other

Appendix E

Metacognition

INSTRUCTIONS: Below are statements describing people's behaviors. Please use the following scale to indicate how accurate each statement reflects **your own** behavioral tendencies.

1	2	3	4
Do not agree	Agree slightly	Agree moderately	Agree very much

1. I am constantly aware of my thinking.
2. I pay close attention to the way my mind works.
3. I think a lot about my thoughts.
4. I constantly examine my thoughts.
5. I monitor my thoughts.
6. I am aware of the way my mind works when I am thinking through a problem.

Wells, A., & Cartwright-Hatton, S. (2004). A short form of the metacognitions questionnaire: Properties of the MCQ-30. *Behaviour Research And Therapy, 42(4)*, 385-396.

Appendix F

Personal Efficacy Beliefs Scale

INSTRUCTIONS: Below are statements reflecting people' ability to do tasks required by their classes. Use the following scale to indicate how accurately each statement describes **your ability** to perform the class-related tasks mentioned below.

1	2	3	4	5
Strongly Disagree	Moderately Disagree	Neutral	Moderately Agree	Strongly Agree

1. I have confidence in my ability to do well in ______ class.
2. There are some tasks required by my _______ class that I cannot do well. (Reversed)
3. When my grades are poor, it is due to my lack of ability. (Reversed)
4. I doubt my ability to do well in my ______ class. (Reversed)
5. I have all the skills needed to perform well in my ______ class.
6. Most people in my class get better grades than I do. (Reversed)
7. I am a great student.
8. My future in school is limited because of my lack of skills. (Reversed)
9. I am very proud of my skills and abilities in school.
10. I feel threatened when others watch me take a test or do homework. (Reversed)

Riggs, M. L., Warka, J., Babasa, B., Betancourt, R., & Hooker, S. (1994). Development and validation of self-efficacy and outcome expectancy scales for job-related applications. *Educational and psychological measurement, 54(3)*, 793-802.

Appendix G

Professionalism

INSTRUCTIONS: Below are statements describing people's behaviors. Please use the following scale to indicate how accurate each statement reflects **your own** behavioral tendencies.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

1. If I realize that I will be late, I contact the appropriate individual at the earliest possible time to inform them.
2. I accept consequences for my actions
3. I follow through with my responsibilities.
4. I am committed to helping others.
5. I would report a medication error even if no one else was aware of the mistake.
6. I address others using appropriate names and titles.

Revised from Chisholm, M. A., Cobb, H., Duke, L., McDuffie, C., & Kennedy, W. K. (2006). Development of an instrument to measure professionalism. *American Journal of Pharmaceutical* Education, 70(4), 1-6

Appendix H

Emotion Control

Some difficulties may have arisen during _____ course. Please choose the response that best describes your reaction to these difficulties.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

When difficulties arose:

1. I did not allow myself to lose my composure.
2. I calmly considered how I could continue the task.
3. I purposely continued to focus myself on the task.
4. I allowed myself to be distracted by worrisome thoughts. (Reversed)
5. I let myself become distracted. (Reversed)
6. I let myself be sidetracked from the task. (Reversed)
7. I was able to focus all my attention on the task.
8. I was able to motivate myself to continue.

Keith, N. (2005). *Self-regulatory processes in error management training* (Doctoral dissertation). Retrieved from Universitätsbibliothek Giessen.