Supplemental Instruction, Calibration, and Self-Efficacy: A Path Model Analysis

Jennifer Grimm

Doctoral Dissertation Proposal, Summer 2018

Old Dominion University

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

To ensure that the United States (U.S.) remains a world leader in STEM education, educators, policymakers, and special interest groups are placing an emphasis on preparing college students for careers in the fields of science, technology, engineering, and mathematics (STEM; Koenig, Schen, Edwards, & Boa, 2012; National Science Foundation, 2011). Regrettably, many students are unable to persist past entry-level courses in STEM fields (Hopper, 2011; Nasr, 2012; Perez, Cromley, & Kaplan, 2014), let alone successfully complete their college degrees (Complete College America, 2014; Kitsantas, Winsler, & Huie, 2008). Increased *access* to higher education does not necessarily translate into academic *success* in entry-level STEM courses (Douglas-Gabriel, 2015; Schudde & Goldrick-Rab, 2016; Smith, 2016). This is due to a variety of factors, including social and economic disparities, which often contribute to a lack of academic preparation prior to college (Douglas-Gabriel, 2015; Pew Research Center, 2014). This lack of preparation leads to poor self-regulated learning (SRL) behaviors, low self-efficacy towards challenging STEM course content, and ultimately insufficient grades to persist into upper-level STEM classes (Bembenutty, 2007; Kitsantas et al., 2008; Perez et al., 2014; Usher, 2009, 2016).

Background

In addition to learning the content necessary to pass entry-level STEM courses, students' self-regulation of their learning activities influences their ability to succeed academically (Schunk & Pajares, 2005). In response, many institutions of higher education have implemented intervention programs to help students review course content and gain the cognitive and metacognitive strategies for success in entry-level STEM courses like general biology (Gattis, 2002; Mack, 2007). One such program is Supplemental Instruction (SI), which has been widely adopted by colleges and universities worldwide (Elam, 2016).

SI is an academic support program that targets historically difficult courses, rather than at-risk students. The goals of SI include increasing students' final course grades, reducing attrition within difficult classes, and improving institutional retention and graduation rates (Arendale, 1997). Instructional faculty of these high-risk courses invite students who have successfully completed their class to serve as SI leaders. These students attend class lectures and follow course readings and assignments. SI leaders also plan weekly, optional, out-of-class group study sessions to provide students with additional opportunities to review class content, work in peer study groups, and develop the SRL behaviors necessary for success in their current and future courses (Arendale, 1997; Elam, 2016; Hurley, Jacobs, & Gilbert, 2006).

Description of the Problem

Numerous studies have demonstrated that SI attendance is correlated with students successfully passing challenging college courses (e.g., Arendale, 1997; Blanc, DeBuhr, & Martin, 1983; Rabitoy, Hoffman, & Person, 2015). However, few studies have used a SRL perspective to examine the SI program's impact on students' self-efficacy or calibration accuracy, which are necessary attributes for college achievement beyond entry-level, SI-supported courses. Self-efficacy is a motivational construct that describes people's convictions about their ability to perform certain tasks (Schunk, 2012). Calibration is a related metacognitive construct that measures how a person's ability to self-monitor and predict their performance matches his or her actual performance (Hacker, Bol, & Keener, 2008). Improvements in the SRL constructs of self-efficacy and calibration accuracy can lead to increased student retention and persistence (Jarvela & Jarvenoja, 2011; Schunk, 1990; Schunk & Pajares, 2005).

It is important to examine connections between SI programs and the SRL constructs of self-efficacy and calibration for two reasons. First, it is practically vital to identify if gains in students' academic success may extend beyond the semester during which students participate in the SI-supported course to help institutions weigh the costs and benefits of a program that requires considerable financial and human resources (Curators of the University of Missouri, 2011). Second, there is value in advancing knowledge on the scarcely explored theoretical connections between SI, self-efficacy, and calibration and the potential mediating effects increases in self-efficacy and calibration may have on students' final course grades.

Purpose Statement

The purpose of this study is to examine the practical connections between a Supplemental Instruction program and the constructs of self-efficacy and calibration. Specifically, I will investigate if students' pre-existing self-efficacy beliefs and calibration accuracy predict their decisions to attend SI sessions throughout the semester. In addition, the study will explore if SI attendance has a direct effect on changes in students' self-efficacy and calibration and subsequent indirect effects on students' final course grades.

Research Questions

Two research questions will guide my study:

1. To what extent do students' self-efficacy beliefs and calibration accuracy at the beginning of a general biology course predict their SI attendance during the semester?
2. Controlling for pretest differences, to what extent does SI attendance predict final calibration accuracy, self-efficacy, and course grades at the end of a general biology course?

Overview of Methodology

I will employ a non-experimental correlational design via a structural equation modeling (SEM) analysis to answer the research questions. The exogenous (or independent) variables included in the hypothesized path model are total SAT score, beginning calibration, and beginning self-efficacy. The endogenous (or dependent) variables are SI attendance, final calibration, final self-efficacy, and final course grade. I will conduct my study with approximately 580 potential participants from an introductory biology course taught by one instructor and supported by the SI program at a large research institution in the Mid-Atlantic region of the United States. Calibration and self-efficacy measures will be administered to participants prior to the first and final course exams, and the other variables will be collected from the SI program and institutional assessment office.

I will apply a path analysis with robust maximum likelihood estimation to answer my research questions using Mplus (v 7.3; Byrne, 2012). Fit statistics recommended by Hu and Bentler (1999) will be used to assess model fit, including chi-square (x^2), Tucker Lewis Index (TLI), root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR). If the hypothesized model is rejected due to its poor fit with the sample data, I will engage in a process known as "model generating" (Byrne, 2012, p. 8) by which I will release one path at a time and analyze the changes in chi-square to determine any statistically significant improvements in the model (Loehlin, 1998). I will describe in detail any such changes in my completed study.

Definition of Terms

A key term used throughout the study is *Supplemental Instruction (SI)*. *SI* is an academic support program that provides students enrolled in historically challenging courses with optional, out-of-class, group review sessions led by student SI leaders (Elam, 2016; Hurley et al., 2006). A major goal of SI programs is to increase students' average course grades and to reduce *DFW rates* within supported classes. *DFW rate* refers to the percentage of students within a course who earn a D or F letter grade or withdraw from the class (Arendale, 1997).

This study uses Zimmerman's (2000, 2002) model of *self-regulated learning (SRL)* as the guiding theoretical framework. Zimmerman's theory of *SRL* stems from Bandura's (1986) social cognitive perspective. According to Zimmerman (2002), "Self-regulation refers to self-generated thoughts, feelings, and actions that are planned and cyclically adapted to the attainment of personal goals" (p. 14). This personal feedback loop consists of three cyclical SRL phases: forethought, performance, and self-reflection. Two constructs found within Zimmerman's model are *self-efficacy* and *calibration*, which are key variables in the present study. *Self-efficacy* is a motivational factor present in Zimmerman's (2002) forethought phase, and it refers to personal convictions held by individuals about their capability to execute behaviors successfully at certain levels (Bandura, 1977; Schunk, 1991; Schunk & Pajares, 2005). *Calibration* is a form of self-monitoring present in all three phases of Zimmerman's SRL theory (Hacker & Bol, in press) and involves measuring how a person's perception of their performance matches his or her actual performance (Hacker et al., 2008).

Delimitations

I have selected several delimitations to guide the scope of my study. First, I have chosen to focus my research study on a general biology course due to its important role in STEM education, its high enrollment numbers, and the control afforded by having one instructor teaching all course sections. In addition, this study examines a Supplemental Instruction program at a four-year research institution because it is an accessible sample and STEM education is important at the institution. I also have decided to limit my study to include only self-efficacy and calibration from Zimmerman's (2002) SRL theory because of clear theoretical connections between both constructs and SI program activities and to simplify my hypothesized path model. To streamline the SEM model further, I have chosen to use total SAT score as a predictor of prior achievement; however, other indices of achievement, including high school GPA, could have been used. Finally, I further chose to limit my study by not including in my path model demographic characteristics such as gender or race/ethnicity.

Significance of the Study

This quantitative study will contribute to SI program and educational psychology research in several ways. First, my research will add to and address the limitations of the few empirical studies that have examined correlations between SI and self-efficacy. This also may be the first study to situate calibration within SI, academic support programs, or help-seeking contexts. In addition, my study will add to the limited empirical literature that has examined how self-efficacy and calibration interact with and influence one another. Finally, this study may produce further insights on the indirect effects of SI attendance (i.e., changes in self-efficacy and/or calibration) on students' final course grades.

Summary

I began this chapter by describing the importance of STEM education in the U.S. and the lack of college students' success in STEM courses caused in part by poor self-regulation of their learning. Many colleges and universities have implemented Supplemental Instruction programs to support students enrolled in challenging entry-level STEM classes. While numerous studies have correlated SI attendance with success in the course, it is important to examine the potential long-term effects of SI attendance on students' SRL constructs of self-efficacy and calibration accuracy. I presented the purposes of my study: (a) to look at how self-efficacy beliefs and calibration may predict students' decisions to attend SI and (b) to explore the effects of SI attendance on students' final self-efficacy, calibration, and course grades. The research questions, methodology, definitions of terms, delimitations, and significance of the study were also presented. In the next chapter, I provide a review of the theoretical and empirical literature on SI, SRL, self-efficacy, calibration, and help seeking.

CHAPTER TWO

REVIEW OF THE LITERATURE

Building on the problem presented in the previous chapter, this review of the literature presents the history, key components, and relevant research related to Supplemental Instruction (SI). I then provide Zimmerman's (2002) theory of self-regulated learning (SRL) which serves as the theoretical basis for the study. I overview SRL, self-efficacy, and calibration, including definitions and key components, theoretical relationships to SI program activities, and relevant research findings, limitations, and implications. Finally, I present help-seeking research literature and conclude with my research questions and summary.

Supplemental Instruction

In this section, I outline the history of the SI program, along with its key components. Then, I present major findings from SI program research along with strengths and limitations of the studies.

History of Supplemental Instruction

Supplemental Instruction (SI) is an academic support model that was developed at the University of Missouri – Kansas City (UMKC) in 1973. The original pilot for the academic support program was for graduate students in the school of dentistry in response to the institution's challenges to retain minority students in its professional schools (Arendale, 1997; Widmar, 1994). The pilot later expanded at UMKC to improve the academic performance and retention of students in high-risk classes in response to first- and second-year student attrition rates of 40 percent.

The SI model was unique in principle because of its focus on high-risk courses, rather than at-risk students (Blanc et al., 1983; Hurley et al., 2006). A collection of prominent learning theories influenced the development of the program model, including cognitivism, constructivism/social constructivism, social interdependence/cooperative learning theories, and critical theory (Bandura, 1977; Freire, 1993; Hurley et al., 2006; Johnson, Johnson, & Holubec, 1994; McGuire, 2006; Zerger, 2008).

After undergoing a rigorous review process by the U.S. Department of Education in 1981, 1985, and 1992, SI became one of the few programs in higher education to receive the coveted status of an Exemplary Educational Program (Martin & Arendale, 1992). SI gained this status because of its three proven claims of effectiveness. First, students who attend SI sessions earn higher final course grade averages than their classmates who do not use the program, even after controlling for ethnicity and prior academic achievement. Second, SI participants succeed at higher rates than non-participants do, despite ethnicity and prior academic achievement. Third, students who participate in SI persist at the institution at higher rates, in terms of reenrollment and graduation, than non-participants do (Martin & Arendale, 1992).

Today, SI programs have been widely adopted by institutions worldwide, with UMKC serving as the International Center for Supplemental Instruction. Through this center, institutions interested in implementing the SI model may send administrators and instructors through the training program for SI supervisors and apply for official SI program certification (UMKC SI, 2018).

Key Components of Supplemental Instruction

The SI model involves several key components that make the academic support program unique, intentional, and effective. This section overviews the major roles of people involved in the implementation of the SI program, courses supported by SI, and factors believed to influence the program’s success.

When presenting the SI model, it is typical to outline the four major roles, or "the four pillars" of SI (Zaritsky & Toce, 2006). These roles include SI supervisors, SI leaders, faculty instructors, and students or college administrators (Hurley et al., 2006; Zaritsky & Toce, 2006). Courses selected for participation in SI programs typically have high rates of students who earn D and F grades and withdrawal from the course (or DFW rates). Typically, SI supports courses with a DFW rate of 30% or above, though this varies by college or university. In addition, institutions typically use SI support for courses that may prevent first- and second-year students from progressing within their major (Hurley et al., 2006).

Blanc et al. (1983) cited six attributes of the SI model that they believe contribute to student success. First, the program is proactive in that students may start benefiting from SI at the beginning of the semester, instead of waiting until it is too late to receive help. Second, the service is connected to a course and its content, rather than general learning skills support. Third, the SI leader's attendance at each class meeting is essential to the program's effectiveness. Fourth, SI is not a remedial program, since it focuses on high-risk courses rather than on struggling students. Fifth, SI sessions involve a lot of student interaction and peer support, leading to positive student academic outcomes. A final unique attribute of SI is the opportunity for the course instructor to receive useful feedback from the SI leader about problems encountered by students (Blanc et al., 1983).

Supplemental Instruction Research

Much research on the SI model has focused on student learning and achievement outcomes, though some researchers also have examined how SI affects student motivation. In this portion of the SI literature review, I outline previous findings related to student academic achievement and motivation outcomes. In addition, I synthesize the methodological strengths as well as limitations and gaps in the literature.

Impact of SI on student learning and achievement. Many SI program research studies have sought to examine the three major claims of the SI model's effectiveness found by the U.S. Department of Education. Again, these three claims include the following: SI participants (a) earn higher final course grade averages; (b) have lower DFW rates; and (c) experience higher rates of reenrollment and graduation (Martin & Arendale, 1992).

SI impact on grades and DFW rates. Many SI studies have resulted in significant correlations between session attendance and increased course grade averages and decreased DFW rates (e.g., Arendale, 1997; Blanc et al., 1983; Grimm & Perez, 2017; Martin & Arendale, 1992; Rabitoy et al., 2015). Many of these studies (e.g., Blanc et al., 1983) distinguished between the SI group and non-SI group based on the number of sessions students attended (e.g., attended 1+ session, 3+ sessions, etc.), while other researchers examined SI attendance frequencies using analysis of variance strategies (e.g., Bruno et al., 2016) or simultaneous multiple regression (Grimm & Perez, 2017; Rabitoy et al, 2015). While most studies examined a single institution, two studies examined the positive impact of SI on course grades and DFW rates at multiple institutions (Arendale, 1997; Martin & Arendale, 1992). In addition, two publications provided a breakdown of SI programs' impact on course grade by examining differences between SI and non-SI participants across top and bottom quartiles determined by institutional admissions standards (Arendale, 1997; Blanc et al., 1983). As noted, results of these studies support the effectiveness of SI on students' performance in supported courses.

A rare instance in which an SI program was not found to have a positive impact on participants' final course grade average was reported by Terrion and Daoust (2012) using a residence study group program, which followed the SI model. While the researchers did find a positive correlation between SI participation and students' likelihood to persist at the institution, there was no correlation between session attendance and final course grades.

SI impact on reenrollment and graduation rates. In addition to Terrion and Daoust's (2012) study, other researchers have examined the impact of SI attendance on students' reenrollment and graduation rates. The home institution for SI (UMKC) was the site for these studies. Arendale (1997) and Martin and Arendale (1992) found that students who attended SI at least one time had higher reenrollment and graduation rates than comparable peers at UMKC. Blanc et al. (1983) also found an increase in retention rates the following semester for students who participated in one or more SI sessions.

SI impact on student motivation and SRL. Outside of traditional academic achievement measures, a few researchers have attempted to examine how SI participation influences students' SRL and/or self-efficacy (e.g., Garcia, 2006; Mack, 2007; Ning & Downing, 2010; Visor, Johnson, & Cole, 1992). These studies have had mixed results, and I discuss them in further detail later in the literature review.

Methodological strengths and limitations of the SI research. A multitude of researchers have sought to examine the impact of students' SI attendance on course grade averages, DFW rates, retention and graduation, and motivation. While all studies have their limitations, there are methodological strengths that are worth examining.

First, several of the studies, though not all, demonstrate that the researchers examined programs that appropriately implemented the SI model (e.g., Dancer, Morrison, & Tarr, 2015; Fayowski & MacMillan, 2008; Terrion & Daoust, 2012). This was apparent through their literature review and methodology sections.

Also, while it can be a limitation that SI program studies are typically non-experimental, a strength is that many researchers accounted for this by including demographic and prior achievement variables to control for the effects of SI attendance on student performance. Control variables used included the following: motivation to attend SI (e.g., Terrion & Daoust, 2012), high school/admissions GPA (e.g., Grimm & Perez, 2017), scores on standardized tests (e.g., Rabitoy et al., 2015), academic rank at the institution (e.g., Gattis, 2002), gender (e.g., Fayowski & MacMillan, 2008), and race/ethnicity (e.g., Mack, 2007).

While strengths exist in the SI research literature, there also are methodological limitations and gaps to address. Specifically, two areas of concern include the necessity for a more consistent way of defining the SI treatment group and a need for more peer-reviewed research on the connections between SI attendance and self-efficacy and SRL.

Inconsistent SI group definitions. First, nearly every study defines the SI treatment group differently. For example, some research studies place students into the SI group if they only have attended one session during the term (Blanc et. al, 1983; Martin & Arendale, 1992), while others require students to have attended two or more sessions (Terrion & Daoust, 2012), three or more sessions (Bowles & Jones, 2003), or five or more sessions (Fayowski & MacMillan, 2008). Thus, there is a great deal of variability in how studies define the SI group. Other researchers have divided participants into more than two groups according to varying levels of attendance and have used analysis of variance or chi-square methods to compare groups. Similarly, these studies have used inconsistent groupings, including: three groups of 0, 1-3, and 4+ sessions (Bruno et. al, 2016; Gattis, 2002); four groups of 0, 1-3, 4-7, and 8+ sessions (Mack, 2007); and five groups of 0, 1-3, 4-7, 8-11, and 12+ sessions (Arendale, 1997).

The International Center for SI's program certification process developed in 2017 establishes a clear set of session attendance groupings that may be useful for future standardization for analysis of variance studies. These groupings examine students who attended 0, 1-4, 5-9, and 10+ sessions throughout the term (UMKC, 2018). However, a continued problem with placing students into SI attendance groups is that the artificial creation of categories may arbitrarily define the number of SI sessions students must attend to reap the program's benefits. For example, the International Center's new categorization (0, 1-4, 5-9, and 10+ sessions) assumes that there is a significant difference between students who attended four sessions versus those who attended five sessions, but that there is no variation between students who attended five sessions and those who attended six sessions. Using linear models of analysis, where SI attendance is a continuous predictor of achievement, can improve understanding of how attending SI relates to achievement (Cohen, 1983).

Rabitoy et al. (2015) used linear multiple regression with SI attendance as a continuous variable and found that SI attendance was a significant positive predictor of increased course grades and cumulative GPA for students enrolled in STEM courses at a Hispanic-serving community college in Southern California. However, the unique nature of the Hispanic-serving community college might limit the generalizability of results to other programs. Grimm and Perez (2017) also used SI attendance as a continuous independent variable in their study. The researchers used longitudinal path modeling to examine the effectiveness of SI attendance on final course grades for students enrolled in two consecutive anatomy and physiology courses. Results indicated that SI attendance in both courses had a significant positive effect on course grades, even after controlling for prior achievement, and that there were indirect effects of attending SI on course grades. More studies like these two examples that use SI attendance as a continuous predictor of achievement can help practitioners better understand how SI session attendance relates to positive academic outcomes.

Need for more theoretically informed research. A second area of concern with the existing literature is that there is a need for more research on SI programs that examines the social cognitive theoretical foundations of the program. I unearthed ten studies on SI programs and student motivation/SRL, and the most recent research on this topic occurred in 2010 (Fisher, 1997; Garcia, 2006; Grier, 2004; Hizer, 2010; Hurley, 2000; Mack, 2007; McGee, 2005; Ning & Downing, 2010; Visor, Johnson, & Cole, 1992; Watters & Ginns, 1997). Now that I have provided an overview of Supplemental Instruction, the next section presents the theoretical framework that informs this study: Zimmerman's model of self-regulated learning.

Self-Regulated Learning

Self-regulated learning theory comes from Bandura's (1986) social cognitive perspective. A commonly used model for describing SRL processes is Zimmerman's (2000, 2002) three-phase model. According to Zimmerman (2002), "Self-regulation refers to self-generated thoughts, feelings, and actions that are planned and cyclically adapted to the attainment of personal goals" (p. 14). This personal feedback loop consists of three cyclical phases: forethought, performance, and self-reflection. Figure 1 presents a visual representation of Zimmerman's (2002) SRL model.

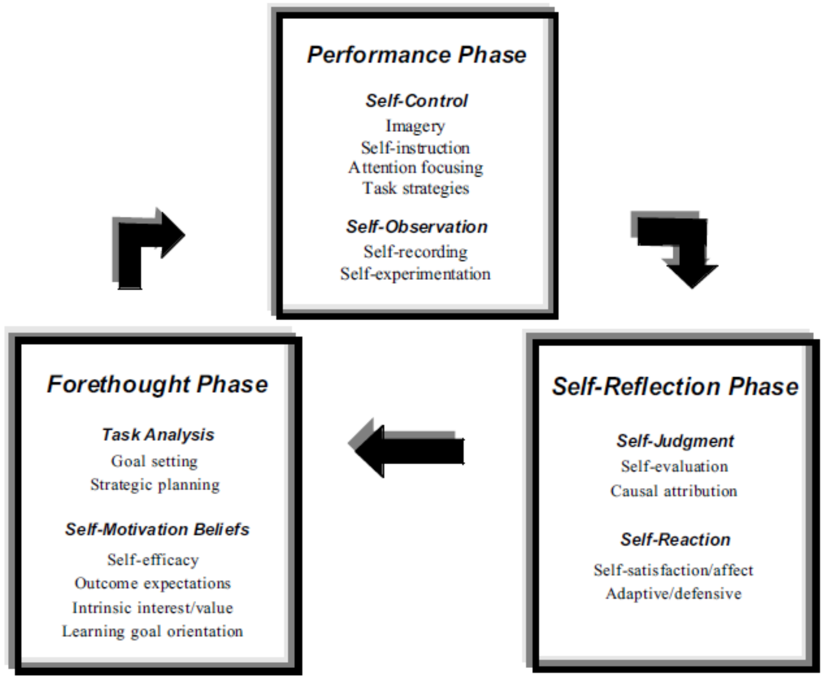

Figure 1. Phases and Subprocesses of Self-Regulation. From Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice, 41* (2), 64-70.

In this section, I describe Bandura's social cognitive theory and detail Zimmerman's (2002) SRL model. Then, I illustrate how SRL behaviors are encouraged during SI sessions. Finally, I outline empirical research on SRL and SI.

Bandura's Social Cognitive Theory

Before detailing Zimmerman's SRL model, I describe Bandura's (1986) social cognitive perspective from which this model is derived. Social cognitive theory (SCT) views humans as agents who are proactively engaged in their own development (Bandura, 1986; Schunk & Pajares, 2005). Bandura's (1986) SCT assumes five basic capabilities that distinguish humans from other lifeforms: vicarious, symbolizing, forethought, self-regulatory, and self-reflective capabilities.

In its most basic format, vicarious learning occurs by observing others modeling behaviors (Bandura, 1986). In addition, people use symbolic processes to help them conceptualize their lived and vicarious experiences into internal guides that they use to direct future actions (Bandura, 1986). An example of a symbolic process is self-efficacy, which involves people's self-evaluations of their capability to perform certain tasks (Schunk, 2012). Like symbolism, forethought is another cognitive capability central to SCT. Once persons create meaningful symbols used to serve as their internal guides, they use this information as they determine how to engage in intentional and purposeful actions. Thus, forethought is heavily engaged in symbolic, as well as self-regulatory, processes (Bandura, 1986).

In addition to vicarious and cognitive capabilities, self-regulatory processes are key tenants of SCT. Self-regulation refers to self-generated thoughts, feelings, and actions, which learners use to set challenging goals for themselves and apply necessary self-regulative strategies to achieve their goals (Schunk, 2012; Zimmerman, Bandura, & Martinez-Pons, 1992). While forethought is heavily present in the early stages of self-regulation, self-reflective capabilities become important after people have determined and pursued their actions (Bandura, 1986). These five capabilities of vicarious experiences, symbolizing, forethought, self-regulation, and self-reflection are reflected in Zimmerman's SRL model.

Zimmerman's Three-Phase Model

In this section, I describe the three phases of Zimmerman's (2002) model of self-regulated learning. Then, I make direct connections between Zimmerman's theoretical model and related SI practices and research.

Forethought phase. The forethought phase of Zimmerman's model consists of task analysis and self-motivation beliefs. During task analysis, learners spend time setting goals, or deciding on their desired learning outcomes or performance. Students also engage in the strategic planning process whereby they identify the methods necessary for reaching their goals (Zimmerman, 2000).

Students' self-motivation beliefs have a strong influence during the forethought phase because self-regulatory behaviors will not occur if people cannot motivate themselves to use them (Zimmerman, 2000). During the forethought phase, learners consider their self-efficacy, or their beliefs about their personal capability to accomplish their goals, along with their outcome expectations, or the personal consequences of learning (Bandura, 1997; Zimmerman, 2002). Furthermore, students are much more likely to be motivated to self-regulate if they have an intrinsic interest and/or see the value in accomplishing their goals. Finally, learners who value the process of learning for its own virtues tend to demonstrate sustained motivation to self-regulate (Zimmerman, 2000, 2002).

Performance phase. During the performance phase, students engage in self-control and self-observation. Self-control involves different strategies for learning content, such as the use of imagery to develop mental pictures and overt or covert self-instruction related to a task. In addition, self-regulated learners improve their concentration through attention-focusing processes, such as setting up an optimal learning environment or ignoring distractions. A final element of self-control involves using task strategies by breaking-down tasks and reorganizing them in meaningful ways (Zimmerman, 2000).

Self-recording and self-experimentation are self-observation strategies used during the performance phase. Students who engage in self-recording keep records of how they used their time to study. In addition, self-regulated learners engage in self-experimentation by trying out different methods for how they spend their time working on a task. For example, a student may self-experiment by studying alone and then with a friend to compare the effectiveness of each study technique (Zimmerman, 2002).

Self-reflection phase. The final phase of Zimmerman's model involves self-reflection through self-judgment and self-reaction. Self-judgment consists of self-evaluation and causal attribution. The first refers to comparing one's own performance against another standard, such as a classmate's or a fixed idea of performance (e.g., earning an A on an assignment). The latter construct, causal attribution, refers to a learner's personal beliefs about the causes of his/her failures. For example, some students will attribute their failure on a math test to a fixed view of their own intelligence, thinking they are simply bad at math (Zimmerman, 2002).

The other part of the self-reflection phase involves self-reaction. The first related construct is self-satisfaction/affect, which refers to people's felt satisfaction or dissatisfaction with their performance. This is important in self-regulation because people tend to act in ways that they believe will lead them to satisfaction and positive feelings, rather than to dissatisfaction and negative affect. Finally, learners make adaptive or defensive inferences to lead them to better forms of performance regulation (i.e., adaptive inferences) or to defensive self-reactions such as task avoidance, procrastination, or helplessness. Thus, these self-reactions have a significant impact on the forethought phase of the cyclical SRL model (Zimmerman, 2000).

Self-Regulated Learning and SI

Self-regulatory process are important influencers of college students' learning and memory (Peverly, Brobst, Graham, & Shaw, 2003) because they help students improve attention, effort, and persistence in coursework for achievement (Jarvela & Jarvenoja, 2011). Thus, there is value in examining the influence SI attendance may have on students' SRL practices. This section examines the theoretical links between SI session activities and Zimmerman's SRL model, as well as relevant research.

SRL and SI sessions. There are clear theoretical connections between Zimmerman's model and the SI model. This is evident in the layout of SI leaders' session plans used to facilitate student learning during sessions. First, like the forethought phase in Zimmerman's model, SI leaders devise an opening activity designed to establish common goals and direction for the session and motivate student attendees. An example of an opening activity is the KWL chart, in which students discuss what they know ("K") and what they want to know by the end of the session ("W"; aka, task analysis). The KWL chart also is commonly used as a closing activity in which students review what they have learned ("L"). Closing session activities like this mirror Zimmerman's third self-reflection phase by providing students with opportunities for self-judgments and self-reactions. Lastly, SI leaders devote most of the session to individual and group learning activities and study strategies, such as the use of imagery and meaningful content organizers that mirror Zimmerman' performance phase (Curators of the University of Missouri, 2011; Zimmerman, 2000, 2002).

SRL and SI research. The clear theoretical connections between SI and SRL have resulted in several studies examining the effect of SI attendance on participants' SRL. Four of the studies used the Motivated Strategies for Learning Questionnaire (MSLQ) to examine effort regulation and resource management (Fisher, 1997; Grier, 2004; Mack, 2007; McGee, 2005), while the other studies used the Learning and Study Strategies Inventory (LASSI; Ning & Downing, 2010) and Study Behaviors Inventory (SBI; Garcia, 2006) to examine students' study behaviors.

First, Grier (2004) looked at the relationship between SI and self-efficacy, outcome expectations, and effort regulation for 43 students in a grant-funded program. Students in this program had the opportunity to participate in SI as a one-credit course in both the fall and spring semesters. The researcher divided students into four groups: (1) non-participants, (2) fall-only participants, (3) spring-only participants, and (4) both fall and spring participants. Students were administered the MSLQ in the summer, fall, and spring. Analyses revealed no significant differences in self-efficacy, outcome expectations, or effort regulation among the four groups. However, generalizability is limited by the special student population examined (i.e., low-income, first generation, and/or nontraditional college students) and SI being offered as a credit-bearing course, as opposed to a voluntary, out-of-class opportunity.

Ning and Downing (2010) used the LASSI to examine various study strategies (e.g., concentration, time management, self-testing, study aids) used by 430 first year undergraduate business students at a university in Hong Kong. Using univariate analyses, the authors found that the 109 students who signed-up for the SI scheme had significantly larger improvements in their pre- and post-test information processing and motivation scores than the 321 students who did not participate in SI.

Garcia (2006) examined the study behaviors of 153 anatomy and physiology students who attended SI sessions. The researcher divided students into a mandatory SI treatment group and a control group that received different interventions of chapter-specific web-based reviews. Garcia (2006) compared both groups' responses to the SBI, and the results showed no statistically significant differences between the groups on any of the three scales: (a)academic self-esteem, (b) time management for the preparation of everyday tasks, and (c) time management for the preparation of long-range academic tasks. Mandatory SI differs from the traditional, voluntary SI model, so this is important to consider when interpreting the results of this study.

Mack (2007) examined the differences in self-regulated learning due to student participation in SI. The researcher administered the MSLQ to 733 students in biology and chemistry courses at a large research university. Mack (2007) divided participants into four groups based on SI attendance: 0, 1-3, 4-7, and 8+ sessions. Results indicated that SI participation did not affect motivation for biology students; however, chemistry students who attended 8+ SI sessions had a positive correlation with motivation on the MSLQ. Furthermore, there were no statistically significant gains for SI participants in the areas of cognition, metacognition, and resource management strategies from the beginning to the end of the semester; though, SI participants in both courses demonstrated resource management at significantly higher rates than non-SI participants in both classes.

McGee (2005) examined the relationship of motivational variables with engagement in SI using the MSLQ as a pretest only for 1,003 students enrolled in biology, chemistry, organic chemistry, horticulture, history, and political science courses supported by SI at large state university. The researcher divided participants into three groups. The first group was of non-participants. The second high-engagement group included students who attended 6+ sessions and received an SI participation score of 2.5+ on a 4.0 scale. The third low-engagement group consisted of participants who attended fewer than six sessions and/or had a participation score below 2.5. McGee (2005) found statistically significant correlations between student participation in SI on 7 of the 11 measured variables, including extrinsic motivation, organization, self-efficacy, effort regulation, control beliefs, peer learning, and help seeking. All correlations were positive with the exceptions of the self-efficacy and control beliefs scales, which had negative correlations with SI participation. The researcher did not administer the MSLQ as a posttest, which means the impact of SI attendance and engagement on students' SRL and motivation is unknown.

Finally, Fisher (1997) sought to determine if participation in SI affects students' motivational orientations and learning strategies. At a large land-grant university, the researcher administered the MSLQ as a posttest to 381 students in three Psychology courses, one of which provided students with the opportunity to attend SI sessions. Results revealed no significant differences between the SI treatment and control groups on 13 of the 15 MSLQ scales, with only significant differences between the groups on the peer-learning and help-seeking scales. However, there were several limitations to this study. First, Fisher (1997) only distributed the MSLQ as a posttest measure, which makes it difficult to know if the groups already differed prior to the SI treatment. Second, students' attendance at SI sessions was restricted to a certain number of SI sessions during the semester, which is not a typical practice of SI programs. Lastly, the author never mentioned how many sessions the SI treatment group attended, which makes it difficult to apply the results to other settings.

In summary, several of the studies were unable to demonstrate or appropriately examine a statistically significant impact of SI attendance on students' SRL capabilities (Fisher, 1997; Garcia, 2006; Grier, 2004; McGee, 2005). Among the studies with statistically significant findings: Ning and Downing (2010) found significant gains for SI participants in the areas of information processing and motivation and Mack (2007) discovered some significant differences in motivation and resource management.

There are four major limitations among the SI and SRL studies. First, two of the studies examined programs that did not follow the SI model (Garcia, 2006; Grier, 2004). Two of the studies also were unable to measure growth from the beginning to the end of the semester due to only administering a pretest (McGee, 2005) or posttest (Fisher, 1997). In addition, as with most SI research studies, there were varying definitions for the SI groups. For example, McGee used three groups based on attendance and engagement levels, while Mack divided participants into four groups based on number of sessions attended.

Lastly, I would argue that these studies attempted to be too broad in scope in looking at the entire construct of SRL, rather than specifying the components of SRL most likely influenced by SI participation. Sitzmann and Ely (2011) propose that there are 16 constructs (e.g., goals, planning, monitoring) found in the various SRL theories. The studies that looked at SI and SRL examined motivation (Garcia, 2006; Grier, 2004; Mack, 2007; McGee, 2005); resource management (Grier, 2004; Mack, 2007); study strategies (Garcia, 2006; Ning & Downing, 2010); planning (Garcia, 2006; McGee, 2005); and cognition, metacognition, and monitoring (Mack, 2007; McGee, 2005). When looking at the impact of SI session attendance on SRL, I have carefully selected for my study the constructs of self-efficacy, a motivational construct in Zimmerman's forethought phase, and calibration, which is arguably present in all three phases of SRL (Hacker & Bol, in press). Next, I discuss the theoretical and practical implications for examining self-efficacy and calibration in my research study.

Self-Efficacy

Self-efficacy is a symbolic process present in the forethought phase of Zimmerman's (2002) model that refers to personal convictions held by individuals about their capability to execute behaviors successfully at certain levels (Bandura, 1977; Schunk, 1991; Schunk & Pajares, 2005). Self-efficacy beliefs influence the choices college students make, including effort expended, length of perseverance when facing obstacles, and resilience in the face of adverse situations (Pajares, 1996, 2002; Schunk, 1990; Schunk & Pajares, 2005). Self-efficacy beliefs are important to students' pursuit of academic tasks because they need to believe they can succeed in those efforts to be motivated to act (Miller et al., 2015). High self-efficacy for college students, when paired with academic competence and SRL behaviors, can lead to higher intellectual performances and more accurate appraisals of abilities (i.e., calibration; Bandura, Barbaranelli, Caprara, & Pastorelli, 1996; Schunk, 2012). The remainder of this section describes how self-efficacy relates to the SI model and empirical research that has examined self-efficacy and SI programs.

Self-Efficacy and SI

Since SI supports students enrolled in challenging first-year college courses like biology (Gattis, 2002; Hurley et al., 2006; Mack, 2007; Zaritsky & Toce, 2006), many SI participants will experience feelings of intimidation or inadequacy when approaching their coursework. Thus, it is important that SI sessions positively influence students' self-efficacy views, while also helping them develop the skills and content knowledge necessary for success in the course (Schunk & Pajares, 2005).

The SI model is a useful tool for positively affecting the four primary sources that influence self-efficacy: mastery experiences, vicarious experiences, social persuasions, and emotional and physiological states (Bandura, 1977; Usher, 2009). First, SI leaders provide mastery experiences by planning sessions that give students hands-on practice and scaffolding the learning process (Hurley et al., 2006). Students undergo vicarious experiences as they engage in group activities and observe modeling by the SI leader and other attendees (Hurley et al., 2006; McGuire, 2006). Positive social persuasions take place as leaders encourage students to participate in activities in a safe, low-risk environment (Hurley et al., 2006). Finally, SI sessions provide a welcoming, non-threatening place to promote positive emotional and physiological states for studying course content (Hurley et al., 2006; McGuire, 2006).

Self-Efficacy and SI Research

Visor, Johnson, and Cole (1992) published the first study to examine motivational factors as they relate to SI. Using the Self-Efficacy Scale, these researchers found that, while results were not statistically significant, SI participants saw a decrease in self-efficacy scores. Visor and his colleagues hypothesized that this was because SI attendees better understood the severity of the challenge and could reevaluate and adjust expectations of their ability, while nonparticipants "remained blissfully ignorant of what it takes to succeed" (p. 17). This theory connects an increase in students' calibration accuracy to a decrease in their self-efficacy, which is one of the primary reasons calibration is the other SRL construct included in this study.

Other studies that have examined students' self-efficacy used a variety of measures, including the Science Teaching Self-Efficacy Belief Instrument (Watters & Ginns, 1997), SBI (Garcia, 2006), MSLQ (Fisher, 1997; Grier, 2004; McGee, 2005), Science Motivation Questionnaire (Hizer, 2010), and a self-designed interview protocol (Hurley, 2000). In the SRL and SI research section, I referenced four of the studies that examined SI and self-efficacy. As a review, Grier (2004) found that there were no significant differences in self-efficacy (or outcome expectations or effort regulation) among SI and non-SI participants. In addition, Garcia's (2006) study resulted in no significant differences between students who received SI support and those who did not on any of the three factors of the SBI, including the academic self-esteem factor, which is related to self-efficacy. McGee (2005) administered a pretest only and found a negative correlation between the self-efficacy scale and SI participation, meaning that students with low self-efficacy were more likely to engage in SI. However, the researcher also discovered that SI participants achieved higher final course grades than their peers who began the semester with higher self-efficacy and did not attend or actively engage in SI sessions. Fisher (1997) used the MSLQ as a posttest only and found significant differences between the SI treatment and control groups on 2 of the 15 scales (peer learning and help seeking), but there were no significant differences between the groups on the self-efficacy scale.

Watters and Ginns (1997) also explored how students' self-efficacy changed because of SI involvement. In their published study, they examined 124 early childhood college students, enrolled in a first-year foundational science course at an Australian university. The researchers divided students into four groups based on their SI participation: (a) no SI attendance, (b) attendance at less than 33% of the offered SI sessions, (c) attendance at 33-66% of the sessions, and (d) attendance at more than 66% of the sessions. Students in the course were administered a pre- and post-test of the Science Teaching Self-Efficacy Belief Instrument. Results showed no significant differences among students who attended and those who did not attend SI. However, the authors administered the instrument once again to students after they took their second semester of the sequential foundational science course, and the high attendance SI group (>66% sessions) saw significant increases in self-efficacy related to the course content the following semester. The authors interpreted their findings to mean that the benefits of SI attendance related to self-efficacy may not be immediate and could potentially take more time to become apparent.

Hizer (2010) examined potential affective benefits, such as increased academic self-efficacy and motivation, for students who participated in SI sessions. The study occurred at a small, public, four-year university in California using a sample of 248 students in biology, chemistry, physics, and psychology courses supported by SI. The researcher administered the Science Motivation Questionnaire as a pre- and post-test to students divided into two groups:

non-participants and those who attended five or more SI sessions. Results showed that students in the SI participation group had initially higher levels of anxiety, but their anxiety decreased over the semester, while non-participants' anxiety levels increased. In addition, Hizer (2010) found that confidence decreased throughout the semester for both groups; however, non-participants had higher levels of initial confidence but ended the semester with lower confidence than students in the SI participation group. This study indicates that SI participation may have a modest positive impact on self-efficacy for students in science courses who attend sessions regularly.

Finally, Hurley (2010) examined the impact of Video SI (VSI) on self-efficacy, self-esteem, test taking anxiety, and students' ability to apply new strategies to other courses. VSI is an adaptation of SI in which courses are videotaped and trained facilitators guide students in processing the material. Hurley implemented a qualitative study in which she conducted and coded student interviews. The researcher found that the course enhanced students' overall motivation. The VSI model differs significantly from traditional SI, and the author used a self-developed questionnaire with no reference to the instrument's validity or reliability, which makes the results of this study less applicable than other SI and self-efficacy research.

In summary, the research on self-efficacy and SI participation is mixed. Some of the studies resulted in no significant differences in self-efficacy between SI and non-SI participants (Fisher, 1997; Garcia, 2006; Grier, 2004; Visor et al., 1992). Studies that produced significant results revealed modest (Hizer, 2010; Hurley, 2010) or delayed (Watters & Ginns, 1997) effects of SI attendance on self-efficacy.

Assessments that are too global can weaken study results, since self-efficacy judgments are task- and domain-specific (Pajares, 1996). Therefore, a limitation of the SI and self-efficacy research is that many studies used instruments that are not task- or domain-specific to measure students' self-efficacy (Fisher, 1997; Hurley, 2000; Grier, 2004; McGee, 2005). In addition, two studies did not administer a pre- and post-test. Fisher (1997) only administered a posttest of the MSLQ, which made it difficult to determine if groups differed significantly prior to the SI intervention, while McGee (2005) administered the MSLQ as a pretest only, which made it impossible to determine if SI participation affected students' self-efficacy. Another limitation is that different authors defined the SI group in varying ways. For example, Visor et al. (1992) used three groupings of students who attended 0, 1-3, or 4+ sessions, while Watters and Ginns (1997) used four groups based on 0%, <33%, 33-66%, or >66% sessions attended. Asking research questions that use SI attendance as a continuous predictor of achievement can improve our understanding of how attending SI relates to increases in self-efficacy and SRL.

Calibration

Like self-efficacy, calibration is present in Zimmerman's (2002) model of self-regulated learning. Calibration involves a form of self-monitoring which measures how a person's perception of their performance matches his or her actual performance (Hacker et al., 2008). Calibration accuracy is important to college students because overconfidence in judging one's abilities may lead to students not studying appropriately for academic tasks and a lowered sense of self-satisfaction toward their courses, while underconfidence may lead to wasted time studying easier concepts (Hacker et al., 2008).

Calibration is a measure of absolute accuracy by which researchers compare people's judgments of their performance with their actual performance. Absolute accuracy is different from relative accuracy, which asks people to compare their performance on one item relative to another. According to Hacker et al. (2008), measuring for absolute accuracy, or calibration, is valuable in educational contexts because it is more reliable and more likely to show stable individual differences. In addition, there are various ways in which calibration may be measured, including global-level judgments (i.e., predicting an overall score on an assessment) and local-level judgments (i.e., item-by-item predictions on a measure; Hacker et al., 2008). In this section, I outline how calibration relates to the SI model, relevant findings in calibration research, and studies that have examined calibration and self-efficacy.

Calibration and SI

It is important to examine the potential impact of SI participation on students' calibration accuracy for three reasons. First, studying calibration judgments and self-efficacy of SI participants allows for the testing of the hypothesis made by Visor et al. (1992) that SI participants saw a decrease in self-efficacy because of their increased ability to evaluate their knowledge of course content. In other words, this study seeks to explain whether a potential decrease in SI participants' sense of self-efficacy is a result of their increased ability to calibrate their anticipated and actual performance on the course's final exam. In addition, no known studies have looked at calibration and help seeking or existing academic support models, let alone specifically at calibration and SI attendance. Finally, since SI session activities influence all three phases of Zimmerman's SRL model, SI attendance has the potential to positively impact both calibration accuracy and academic performance (Hacker & Bol, in press).

Calibration Research

While calibration research has not focused on SI or related academic support programs, research from related settings can shed light on how SI attendance may influence students' calibrations. This section outlines consistent findings in calibration research and findings of interventions that target all three phases of Zimmerman's SRL model.

Consistent findings. A few findings appear to be consistent in calibration research. First, high-achieving students tend to be more accurate in their predictions than low-achieving students, and low achievers are often overconfident in their judgments, while high achievers tend to underpredict their performance (e.g., Bol & Hacker, 2001; Flannelly, 2001; Nietfeld, Cao, & Osborne, 2006; Shaughnessy, 1979). Since SI research demonstrates that students who attend SI perform better than their peers, one could surmise that SI participants may make more accurate confidence judgments than students who do not attend SI.

A second common finding is that people's confidence judgments typically remain consistent over time, regardless of their performance (Hacker, Bol, Horgan, & Rakow, 2000; Nietfeld et al., 2006). This finding, contrary to its predecessor, may indicate that any academic intervention (e.g., SI) may not be able to influence students' calibration accuracy.

A last consistent finding is that postdiction judgments (made after an assessment) tend to be more accurate than predictions (made before an assessment; Bol, Riggs, Hacker, Dickerson, & Nunnery, 2010; Dinsmore & Parkinson, 2013; List & Alexander, 2015). For this reason, it is particularly important to assess the impact of SI attendance on students' predictions, since they tend to be less accurate than postdictions.

Interventions targeting all three SRL phases. While some findings remain generally consistent in calibration research, interventions developed to increase calibration accuracy and academic performance have had mixed results. For example, some studies have demonstrated that certain educational interventions increased both calibration accuracy and academic performance (e.g., Bol, Hacker, Walck, & Nunnery, 2012; Morrison, Bol, Ross, & Watson, 2015), while other studies improved calibration with no effects on academic performance (e.g., Huff & Nietfeld, 2009; Reid, Morrison, & Bol, 2016). Hacker and Bol (in press) argue that calibration has implications in all three phases of Zimmerman's cyclical model and that interventions that target all three phases (e.g., SI) will be more successful at improving calibration accuracy and academic performance (Bol et al., 2012; DiGiacomo & Chen, 2016; Gutierrez & Schraw, 2015).

Specifically, Bol et al. (2012) staged a 2 x 2 factorial quasi-experimental design intervention to investigate the calibration accuracy and achievement of 82 high school biology students who used guidelines within group or individual settings. The process of having students predict exam grades and plan review activities activated the forethought phase. Then, the performance phase was involved via use of guidelines and group-led discussions. Finally, making postdictions triggered the self-reflection phase. Participants who received guidelines within group settings had better calibration accuracy and higher exam scores than their peers who were exposed to only one or neither of the interventions.

DiGiacomo and Chen (2016) used an intervention that targeted calibration practices across all three phases of Zimmerman's SRL model. The researchers provided structured, guided questions to 30 sixth and seventh grade students in randomly assigned treatment or delayed treatment control groups to help them review the material and make pre- and post-diction judgments. Then, students received feedback and completed self-reflective worksheets. Study results demonstrated that students in the treatment group, when compared with the control group, had significantly higher math performance, as well as increased pre- and post-dictive calibration accuracy.

Gutierrez and Schraw (2015) also incorporated all three phases of Zimmerman's model. In their study, 107 individuals in randomly assigned treatment groups received cognitive strategies instruction related to calibration accuracy (performance phase), financial incentives for high performance (forethought phase), or both. Participants also made confidence judgments after completing items (self-reflection phase). The researchers found significant effects for the strategy training on performance, confidence, and calibration accuracy, and incentives further improved performance and calibration accuracy.

While none of the described interventions exactly mirrors the Supplemental Instruction model, there are similarities in SI leaders' session plans involving opening, review, and closing activities and Zimmerman's three phases of forethought, performance, and self-reflection (Curators of the University of Missouri, 2011; Zimmerman, 2002). This indicates that SI participants may potentially see improved calibration accuracy and exam scores from their participation in the educational intervention. The study by Bol et al. (2012) has especially significant implications, as students who were provided with guidelines in group settings had the highest calibration and performance among all the groups, which is important because SI sessions also take place within a group setting. However, it should be noted that, unlike the Bol et al. (2012), DiGiacomo and Chen (2016), and Gutierrez and Schraw (2015) studies, the control gained by random assignment will not be possible in the current context of SI support for biology students due to the voluntary nature of SI.

Calibration and Self-Efficacy Research

An area of calibration research that has garnered little attention is the exploration of the interplay between individuals' calibration accuracy and self-efficacy. An important feature of students' self-regulation is their ability to calibrate between their confidence of knowing and actual performance (Bandura, 1986), which is why understanding individuals' contributing motivational factors is a key component in the study of self-regulation. Specifically, self-efficacy, calibration, self-regulation, and motivation are all related concepts (Bembenutty, 2009).

A simplified way of looking at calibration and self-efficacy is that they both involve self-confidence judgments, but calibration is metacognitive in nature, while self-efficacy examines affective or motivational influences.

Chen (2003) studied the calibration and self-efficacy beliefs of seventh grade math students, specifically focusing on whether their calibration was a significant feature of their self-efficacy beliefs. The researcher used a path analysis to examine the interplay of five separate measures, including a math performance test, a math self-efficacy scale, a math effort judgment scale, a self-evaluation scale, and previous math achievement. The results indicated that calibration accuracy had a significant direct effect on students' math performance, as well as an indirect effect on math performance via its significant effect on students' math self-efficacy judgments. Furthermore, self-efficacy played a direct role in predicting students' math performance and this impact was much greater when they also possessed the underlying math skills. Chen (2003) also discovered that students' pre-performance self-efficacy beliefs regarding their math capability had a much larger impact on their post-performance self-evaluations than their math performance, which indicates stable self-views among students, regardless of actual performance. A final notable finding from Chen's (2003) study of seventh grade math students was that participants generally overestimated their math capabilities, but there was no relationship between their inaccuracies and the strength of their self-efficacy beliefs.

Nietfeld et al. (2006) explored how college students' changes in monitoring over the course of a semester affected changes in their self-efficacy from the beginning to the end of the semester. Using a repeated-measures design, 84 undergraduate students in an educational psychology survey course completed weekly monitoring worksheets throughout the term. The researchers provided students with an educational psychology self-efficacy inventory as a pre- and post-test and found a significant effect of average monitoring accuracy on self-efficacy; however, there was no significant effect of the change in calibration accuracy from the beginning to the end of the semester on students' self-efficacy. The researchers asserted that their study demonstrates that even modest metacognitive monitoring interventions can significantly improve students' calibration, performance, and self-efficacy.

In a non-educational setting, Hong, Hwang, Tai, and Chen (2014) studied participants' use of an iPhone application for English vocabulary practice to explain smartphone self-efficacy (SSE) in relation to their judgments of over-confidence (JOOC). Using a path model, the researchers found SSE to be a negative predictor of participants' overconfidence, indicating that those with higher self-efficacy were less likely to over-predict their performance and thus had greater calibration accuracy.

Taken together, these studies indicate a positive significant relationship between individuals' calibration accuracy and self-efficacy (Chen, 2003; Hong et al., 2014; Nietfeld et al., 2006). In addition, Chen's (2003) finding that students' beliefs are likely to remain stable over time regardless of actual performance likely explains the finding in the study by Nietfeld et al. (2006) that average calibration accuracy was a significant positive predictor of self-efficacy while change in calibration accuracy was not a significant predictor of self-efficacy. Finally, the assertions by Nietfeld et al. (2006) that modest metacognitive monitoring interventions can improve students' calibration accuracy, self-efficacy, and academic performance identifies the potential benefits students may experience through participation in SI sessions.

Help Seeking

Since SI attendance will occur only if students choose to participate, it is essential to understand findings from the help-seeking research. Karabenick and Berger (2013) define help seeking as "the process of seeking assistance from other individuals or other sources that facilitate accomplishing desired goals, which in an academic context may consist of completing assignments or satisfactory test performance" (p. 238). I begin this section with two prominent themes in the help-seeking literature: a lack of help-seeking behaviors among students and the two types of help sought by students. Then, direct theoretical connections are made between help seeking and SRL, self-efficacy, and calibration.

Prominent Themes in the Help-Seeking Literature

One major finding in studies of student help seeking is that often students do not seek the help they require to be academically successful. In a study of college students from three diverse institutions, Karabenick and Knapp (1991) discovered that the students who were most in need of help, due to poor self-regulation and study skills, were the least likely to seek help. The researchers suggest several possible reasons why students who most need help were unlikely to seek it out, including: hopelessness, feeling threatened to display their ignorance to others, and a general lack of help-seeking skills or awareness of resources. These reasons for students not seeking help can be problematic with a voluntary academic support program like SI in which students may not take advantage of the help this service provides. This is a reason why it is important to study the metacognitive and motivational factors (e.g., calibration and self-efficacy) that may influence students' help-seeking behaviors.

Another prominent theme in the help-seeking literature describes the two types of help seeking in which students engage: executive help seeking and adaptive help seeking. Executive help seeking occurs when students enlist the help of others to decrease the amount of effort required to complete a task (e.g., getting answers to a problem; Karabenick & Knapp, 1991). Executive help seeking is contrasted with adaptive help seeking whereby students seek the minimum amount of help needed to achieve a task independently. This could involve asking for an explanation or hints rather than direct help with resolving a question (Karabenick & Knapp, 1991). Adaptive help seeking is a self-regulated learning strategy that is goal-directed and intentional in action, and it is different from other SRL strategies that students may employ because of its social origins (Newman, 2008).

Student participants with adaptive help-seeking orientations are ideal attendees of SI sessions. Karabenick and Berger (2013) recommend that interventions designed to promote adaptive help seeking among college students require a comprehensive approach that addresses several competencies and resources, including cognitive, affective-emotional, contextual, and social entities. Interventions achieve the cognitive competency by helping students become aware of their need for help. SI promotes cognitive competency through SI leaders' first-day introduction speeches in which they describe the difficulty of the class material and the importance of mastering the material before moving to upper-level courses (Curators of the University of Missouri, 2011).

Affective-emotional components are also important in developing adaptive help-seeking behaviors. This competency is achieved during SI sessions because they typically promote positive emotional experiences for students as they engage in non-threatening, peer-to-peer environments (Hurley et al., 2006; McGuire, 2006).

In addition, the promotion of adaptive help seeking must be contextual. For example, teachers may establish and explain classroom norms for seeking help. Again, SI leaders' first-day introductions and verbal encouragements from the course instructor prompt students to participate in SI sessions as a way of receiving help with the course (Curators of the University of Missouri, 2011).

A final component to promoting adaptive help seeking is social competence, which involves the social skills required to ask for help. SI helps reduce the challenges students may experience asking for help by providing them with a designated time, place, and environment in which they can show up to review course material and ask questions (Curators of the University of Missouri, 2011). This differs from other forms of help seeking, such as setting up an appointment with the professor. In summary, the SI model addresses the various competencies and resources Karabenick and Berger (2013) describe as necessary for interventions to promote adaptive help seeking, which should translate into adaptive help-seeking behaviors among SI participants. Next, I relate the help-seeking literature to SRL, self-efficacy, and calibration.

SRL and Help Seeking

Self-regulated learning and help seeking are closely related. Karabenick and Berger (2013) make clear connections between the stages of the help-seeking process and Zimmerman's phases and processes involved in self-regulation (see Table 1 below).

Table 1

Help-Seeking Process and Zimmerman's SRL Phases

Stages of the Help-Seeking Process		Processes of SRL	SRL Phase
1	Determine whether there is a problem	Task analysis	Forethought
2	Determine whether help is needed/ wanted		
3	Decide whether to seek help	Strategic planning	
4	Decide on the type of help (goal)		
5	Decide on whom to ask		
6	Solicit help	Self-control	Performance
7	Obtain help		
8a	Process the help received – judge or evaluate it	Self-judgment: self-evaluation	Self-Reflection
8b	Process the help received – react to it	Self-reaction: self-satisfaction and adaptive inference	

From: Karabenick, S. A., & Berger, J. L. (2013). Help seeking as a self-regulated learning strategy. In H. Bembenutty & T. J. Cleary (Eds.), *Self-regulated learning across diverse disciplines: A tribute to Barry J. Zimmerman* (pp. 237-261). Information Age Publishing, Inc.

The forethought phase is involved in the initial five help-seeking steps. The first and second steps, which involve determining if there is a problem and determining whether help is needed or wanted, are components of the task analysis, or more specifically the goal setting, portion of the forethought phase. Then, strategic planning is devised in the following three steps of the help-seeking process by which students decide on whether to seek help, the type of help desired, and whom to ask (Karabenick & Berger, 2013).

Students engage in Zimmerman's performance phase as they demonstrate the self-control required for steps 6 and 7 of the help-seeking process: Solicit help and obtain help. Finally, the last step of help seeking is to process the help received by judging or evaluating it and reacting to it. This last step mirrors Zimmerman's self-reflection phase via self-judgments and self-reactions (Karabenick & Berger, 2013).

This model demonstrates that all three SRL phases must influence students' help-seeking behaviors for them to identify, seek, and reflect on help received (Karabenick & Berger, 2013).

The calibration literature also reveals that interventions that target student monitoring during all three phases of Zimmerman's SRL model tend to be more effective in positively influencing students' calibration accuracy and academic performance (Hacker & Bol, in press). Thus, it is important that SI supports students during all stages of the help-seeking and SRL processes. In addition, the SI model encourages students to identify the need for help and subsequently engage in adaptive help-seeking behaviors, based on the necessary intervention competencies outlined by Karabenick and Berger (2013). Furthermore, it seems that the activities engaged in during SI sessions should allow students to more accurately calibrate their knowledge and skills with their subsequent academic performance. This is particularly characteristic of the beginning activities that help students identify what they already do or do not know and the closing activities that involve self-reflective practices. The process of closing the loop in Zimmerman's three-phase, cyclical model should thus encourage students to continually identify their need (or lack thereof) for additional help with reviewing the course material and influence their decisions to continue to engage in (or not continue to attend) SI sessions.

Self-Efficacy, Calibration, and Help Seeking

For the purposes of the present study, I am interested in the motivational and metacognitive attributes of students that will be most influenced by the elements of the SI model that prompt students to seek help. While achievement goal theory is the motivational theory most commonly associated with help seeking (Karabenick, 2003; Karabenick & Berger, 2013), there are also connections between students' self-efficacy and their help-seeking behaviors. Newman (2008) related help seeking and self-efficacy to students' adaptive and non-adaptive behaviors, as well as to students' performance and mastery goal orientations (See Table 2 for a simplified version of Newman's model).

Table 2

Help Seeking & Self-Efficacy

Is Help Necessary?	Action: Seek Help	Action: Do Not Seek Help
Yes	Quadrant I – Adaptive Goals: Mastery Self-beliefs: High self-efficacy	Quadrant II – Nonadaptive Goals: Performance-avoidance Self-beliefs: Low self-efficacy
No	Quadrant III – Nonadaptive Goals: Performance-approach Self-beliefs: Low self-efficacy	Quadrant IV – Adaptive Goals: Mastery Self-beliefs: High self-efficacy

From: Newman, R. S. (2008). The motivational role of adaptive help seeking in self-regulated learning. In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 315-337). New York, NY: Lawrence Erlbaum Associates.

In Newman's theoretical article, he describes the basic process students go through in help seeking. First, they ask if help is necessary, and then they determine if they will act by seeking help or not. Students who exhibit adaptive behaviors identify that help is necessary and seek it and or identify that they do not need help and do not seek it. These students tend to have a mastery goal orientation and high self-efficacy. Conversely, students with performance goal orientations are more likely to have low self-efficacy and engage in nonadaptive behaviors by identifying that they need help and not seeking it or by seeking help even when they do not have the need. Thus, students with high self-efficacy are more likely to engage in positive, constructive help-seeking behaviors.

Newman's (2008) model is useful for drawing tentative conclusions about the influence of self-efficacy and calibration on students' help-seeking behaviors. First, it appears that self-efficacy will influence students' calibration accuracy. This is demonstrated by the adaptive help seekers who identify their need for help and seek it out, as well as those who identify that they do not need the help and do not seek it out. In other words, students who can more accurately calibrate their need for help are likely to have higher self-efficacy. Conversely, it seems that students who have low self-efficacy may not seek needed help, even when they have identified they need the assistance, while others with low self-efficacy may seek out the help when they do not require the additional support.

Justification for Study

When considering a well-established higher education academic support program like Supplemental Instruction, there is value in knowing that there are strong correlations between SI attendance and students' course grades and passing rates (e.g., Arendale, 1997; Rabitoy et al., 2015). However, success solely in the entry-level STEM courses supported by SI is not enough to sustain students' success throughout their entire academic careers, especially within challenging STEM majors. Existing literature demonstrates that there are many lasting benefits experienced by college students with a strong sense of self-efficacy towards their courses and the ability to accurately calibrate or monitor their performance in their classes. These benefits include an increase in expended effort and resilience through obstacles (Schunk & Pajares, 2005) and appropriate allocation of time spent studying relevant material for success in the course (Hacker et al., 2008). Thus, while the influence of SI attendance on individual course grades is helpful, the potential of the SI model to influence students' overarching metacognitive and affective attributes could have much larger implications.

The influence of students' self-efficacy and calibration abilities are important to examine because there are clear theoretical connections between these constructs and the SI model. SI participation has the potential to affect the four sources that influence self-efficacy (Bandura, 1977; Usher, 2009). Furthermore, SI sessions provide students with informal opportunities to calibrate their perceived knowledge of the course material with their actual knowledge through activities that align with the three phases of Zimmerman's SRL model (Zimmerman, 2000, 2002). In addition, the help-seeking literature acknowledges that different barriers may prevent students from choosing to seek help (Karabenick & Berger, 2013; Karabenick & Knapp, 1991), which makes it important to examine the influence of students' self-efficacy and calibration on their decisions to attend SI sessions.

In addition to examining the potential influence of SI participation on students' self-efficacy and calibration, there is also merit from a theoretical perspective in further studying the complex interplay between self-efficacy and calibration. Calibration accuracy is related to students' metacognitive views on their ability to predict their performance on an assessment (Hacker et al., 2008), while self-efficacy measures students' beliefs on their ability to complete specific tasks (Bandura, 1977). Both constructs are closely related; however, calibration is a metacognitive factor, while self-efficacy addresses affect or motivation. Examining both constructs in the same setting can build upon previous research (e.g., Nietfeld et al., 2006) to help educational researchers have a better understanding of how these cognitive and affective functions interact with each other.

Additional research is needed on how calibration influences help seeking, as well as how academic support programs like SI may influence changes in calibration. In addition, while some studies have examined SI and self-efficacy, there are several limitations to these studies. First, there is inconsistency in the definitions used to distinguish the SI treatment group. In addition, most studies use instruments that are too global in nature to measure the self-efficacy, and some researchers did not administer both pre- and post-tests to measure changes in self-efficacy. Finally, some studies examined programs that do not faithfully administer the SI model.

In summary, there are several reasons why this study is important. First, there is a need to study the potential influence of the SI model's impact on college students' metacognition and motivation that may influence them beyond a single course. There also are strong theoretical connections between SI participation and self-efficacy and calibration, which are constructs with positive, long-term academic outcomes. In addition, there are theoretical interests in examining the related yet distinctive constructs of self-efficacy and calibration within the same study. Finally, there are significant gaps and concerning limitations to address within the existing research.

Research Questions

To add to the body of research on SI programs, self-efficacy, and calibration, I have developed two research questions, including:

1. To what extent do students' self-efficacy beliefs and calibration accuracy at the beginning of a general biology course predict their SI attendance during the semester?
2. Controlling for pretest differences, to what extent does SI attendance predict final calibration accuracy, self-efficacy, and course grades at the end of a general biology course?

Summary

Supplemental Instruction is an academic support program known for helping students succeed in challenging college STEM courses. While this is valuable for the promotion of student success in entry-level STEM classes, it is less clear is if the SI model's influence on student course grades also is associated with broader implications for students' self-regulated learning behaviors that may continue with them throughout college.

This review of the literature has provided theoretical connections between the SI model and SRL strategies, specifically focusing on self-efficacy and calibration. In addition, this chapter has provided an overview of the findings in the empirical research on the interactions between SI, SRL, self-efficacy, and calibration. Specifically, most studies on SRL and SI revealed statistically non-significant results when examining the impact of SI attendance on students' SRL behaviors; though, some researchers did unearth significant gains for SI participants in the areas of motivation, information processing, and resource management. Similarly, several of the studies on self-efficacy and SI resulted in no statistically significant differences between SI and non-SI participants; though, a few of the studies demonstrated modest or delayed effects of SI attendance on self-efficacy. A review of the empirical literature also revealed no research on calibration and SI; however, it demonstrated the potential positive effects that an intervention that influences all three stages of Zimmerman's SRL model (like SI) may have on calibration accuracy and academic outcomes. This chapter also has outlined significant gaps in the empirical research on the interactions between SI, self-efficacy, calibration, and academic outcomes, as well as key findings from the help-seeking literature. The next chapter describes the methodology I will use to answer the research questions derived from this review of the existing literature.

CHAPTER THREE

METHODOLOGY

The previous chapter analyzed the existing literature on Supplemental Instruction, self-efficacy, and calibration, including research findings, strengths, limitations, and gaps that led to the present study. The current chapter describes the methodology I will use to address my research questions and hypotheses, including the study design, participants and context, measures, procedure, data analysis, and foreseeable limitations.

Research Questions

Again, the following research questions will guide the present study:

1. To what extent do students' self-efficacy beliefs and calibration accuracy at the beginning of a general biology course predict their SI attendance during the semester?
2. Controlling for pretest differences, to what extent does SI attendance predict final calibration accuracy, self-efficacy, and course grades at the end of a general biology course?

Hypotheses

The first research question addresses the influence of pre-existing self-efficacy beliefs and calibration capabilities on students' SI session attendance. Previous studies on students' initial self-efficacy and their SI attendance indicate that students with lower self-efficacy are more likely to participate in SI (Hizer, 2010; McGee, 2005). However, in the help-seeking literature, Newman (2008) suggests from a theoretical perspective that students with high self-efficacy and the ability to predict their need for help will participate in an academic support intervention, like SI, if they determine it is needed (or they will not participate if they do not determine that it is needed). Thus, it is unknown if self-efficacy will be positively or negatively correlated with SI attendance. In addition, since no existing research looks at calibration and help-seeking behaviors, it is unknown if calibration accuracy will predict students' SI attendance.

The second research question examines students' SI attendance throughout the semester and its potential correlations with final calibration accuracy, self-efficacy, and course grade. First, to examine SI attendance and calibration, I expect that SI attendance will predict a positive change from beginning to final calibration accuracy; however, it is unclear if SI attendance will predict a positive change from beginning to final self-efficacy due to the potential interactions between final calibration and self-efficacy. Since no one has studied SI participation and calibration, theoretical connections are useful for this hypothesis. Hacker and Bol (in press) argue that interventions that target all three phases of Zimmerman's SRL model are more likely to improve calibration accuracy and academic performance. Since the SI model also follows the three phases of SRL, I expect a positive correlation between SI participation and final calibration accuracy. In addition, calibration research demonstrates that high-achieving students tend to be more accurate in their predictions (e.g., Hacker et al. 2008). Since students who attend SI tend to perform better in the course (e.g., Grimm & Perez, 2017; Rabitoy et al., 2015), it seems likely that those who participate in SI will perform better and have better final calibration accuracy than their peers from the course.

The second research question also examines how SI attendance and calibration may predict changes in final self-efficacy. The effect SI attendance will have on final self-efficacy is not clear. It is likely that all students in the course will have lower self-efficacy by the end of the semester, but the decrease in self-efficacy will be less dramatic for frequent SI participants (Hizer, 2010). However, it also is possible that the effect of SI attendance on final self-efficacy will not be detectable by the end of the semester (Fisher, 1997; Garcia, 2006; Grier, 2004;

Watters & Ginns, 1997). Changes in final calibration may also affect the potential correlation between SI attendance and final self-efficacy, as Visor and his colleagues (1992) surmised that frequent SI participants had lowered self-efficacy because of their increased awareness of, or ability to calibrate, what they did and did not know.

Research question two also asks about the direct and indirect effects SI attendance could have on final course grade. Previous SI research indicates that SI attendance will predict an increase in students' final course grades (e.g., Grimm & Perez, 2017; Rabitoy et al., 2015). What is less evident is whether the increase in final course grades will be partially attributable to the indirect effects of SI attendance on increases in final calibration accuracy and/or self-efficacy. Nietfeld et al. (2006) suggest that even modest metacognitive monitoring interventions like SI can improve students' calibration accuracy, self-efficacy, and academic performance, which is why I predict that increases in calibration and self-efficacy will have indirect effects of SI attendance on final course grade.

Research Design

I will employ a non-experimental correlational design via a structural equation modeling (SEM) analysis to address the research questions. SEM is a statistical methodology that uses a hypothesis-testing approach on a phenomenon to represent causal processes among multiple variables (Byrne, 2012). This method involves pictorially modeled structural relations that represent a series of regression equations tested for adequate goodness-of-fit (Byrne, 2012). Kline (2016) defines SEM as a causal inference method that uses three inputs to generate three outputs. The three inputs, which are present in the current study, include: (1) qualitative causal hypotheses based on theory or empirical findings, (2) questions about causal relations among study variables, and (3) data that are often used from non-experimental designs. The three outputs to be generated in the SEM include: (1) numeric estimates of model parameters for the hypothesized effects, (2) a set of logical implications of the model, and (3) the degree to which the data support the testable implications of the model.

SEM is useful for answering the study's research questions because it involves analyzing data for inferential purposes and estimating the direct and indirect effects of variables (Byrne, 2012). Thus, an SEM model will allow for the identification of potential direct and indirect effects of SI attendance on students' final course grade. Specifically, I will conduct a path model analysis, which, according to Kline (2016), is a commonly used model in SEM. A path model is useful for the present study because each variable can be described with a single measure (e.g., beginning self-efficacy), and the sample size may not be large enough to warrant a full SEM (Kline, 2016). Figure 2 depicts my hypothesized path model.

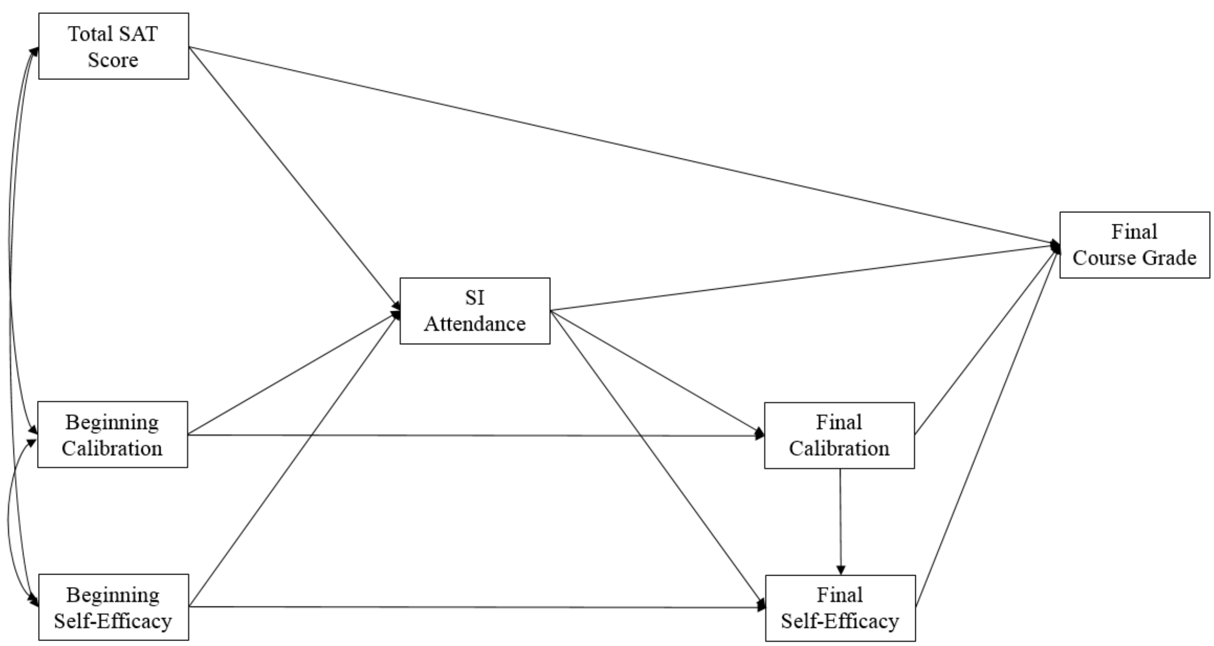

Figure 2. Hypothesized path model to be tested to determine relationships among total SAT score, beginning calibration and self-efficacy, SI attendance, final calibration and self-efficacy, and final course grade.

Participants

The SI program at the institution used for this study supports students in an introductory biology course for science majors each fall semester. While I will not collect data for my study until fall 2018, participant information from the fall 2017 semester is available to me as an administrator of the SI program. One instructor teaches three sections of this course, and 579 students were enrolled in the course in fall 2017. Among these students, 233 (40%) were Caucasian; 195 (34%) were African-American; 51 (9%) were Hispanic/Latino; 34 (6%) reported two or more races; 32 (6%) were Asian; 9 (2%) were Native Hawaiian/Pacific Islander, Native American, or non-resident/aliens; and 25 (4%) declined to report their race. The class also consisted of 385 (66%) females and 194 (34%) males, with the two most popular student majors being biology (n=248; 43%) and physical education (n=117; 20%). The course consisted of 246 (42%) freshmen, 204 (35%) sophomores, 84 (15%) juniors, 37 (6%) seniors, and 8 (1%) masters or unclassified students. In fall 2017, 170 (29%) students attended at least one SI session during the semester.

Hancock and Mueller (2010) recommend having a minimum of five participants per parameter in an SEM model to obtain trustworthy maximum likelihood (ML), while Kline (2016) recommends at least a 10:1 sample-size-to-parameters ratio. The number of parameters in a hypothesized model can be challenging to calculate, but a general guideline is $p \leq K(K + 1)/2$, where K is the number of observed variables in the path model (StataCorp LLC, 2018). Thus, the hypothesized path model (see Figure 2) will have up to 28 parameters (7*8/2), which means that my study ideally should have a minimum of 140-280 participants (Hancock & Mueller, 2010; Kline, 2016). This means that approximately 24 to 48 percent of students in the course will need to complete both the pre- and post-test measures to reach this minimum participant goal. To attempt to reach this number, the course instructor will offer extra credit for completion of the surveys, and I will provide participants with the opportunity to enter their names into a drawing for Amazon gift cards.

University Context

The SI program serves courses at a large research institution in the Mid-Atlantic region of the United States with nearly 20,000 undergraduate students and over 4,500 graduate students who represent a diverse community in terms of race and ethnicity, country of origin, traditional first-year and transfer students, and other factors. Specifically, the university is 56% female and 44% male, and the race/ethnicity of the student population is 47.7% white, 27.4% African American, 7.9 % Hispanic, 4.4% Asian, and 12.6% other/multiple categories. This institution was selected because it provides a convenience sample, and its diverse student population mirrors the demographics of similar institutions. I also chose to conduct my research at this university because it has achieved SI program certification recognition by the International Center for SI at the University of Missouri-Kansas City.

Supplemental Instruction Program

As a certified SI program, the International Center has verified that the institution in the present study has successfully adopted what is referred to as the “Core Four:” (1) training by the International Center, (2) SI leader training and support, (3) a strong focus on planning for sessions, and (4) class attendance and data collection and reporting (UMKC, 2018). By providing evidence of achievement in these areas, the SI program in this study has demonstrated that it closely follows the SI model.

Two trained SI leaders will support three sections of the general biology course that are taught by the same instructor. The SI leaders will be trained on the SI model, including the use of key facilitation strategies and the development and implementation of SI session plans. The SI supervisor will observe both leaders during their sessions throughout the semester to ensure they are following the SI model. After session observations, the SI leaders will receive a completed feedback form and meet with the SI supervisor to discuss their strengths and areas for improvement. Beginning the second week of classes, both leaders will host a combined 6 one-hour SI sessions each week for approximately 78 session attendance opportunities over the course of 13 weeks.

Measures

This section describes the measures to be used in the study. I will administer to participants two scales as pre- and post-tests to measure beginning calibration and self-efficacy early in the semester and final calibration and self-efficacy at the end of the semester. In addition, I will collect SI attendance data from the SI program and will request other information from the institutional assessment office and course instructor.

The path model includes three exogenous (or independent) variables and four endogenous (or dependent) variables. Exogenous variables cause fluctuations in other variables in the path model and are influenced by factors that are external to the model (Byrne, 2012). The exogenous variables in the current study are total SAT score, beginning calibration, and beginning self-efficacy. Endogenous variables are influenced by the exogenous variables, either directly or indirectly (Byrne, 2012). SI attendance, final calibration, final self-efficacy, and final course grade are the endogenous variables in the path model.

Calibration

Calibration describes how well one can judge their performance on a task (Bol et al., 2010). In my study, the tasks are the first and final exams taken by students, which will be used to measure the beginning calibration and final calibration variables.

Beginning calibration. Beginning calibration is an exogenous variable within the hypothesized path model that may cause fluctuations in SI attendance and final calibration. To measure beginning calibration, I will ask students on the pretest survey prior to their first exam: "On a scale of 0-100%, predict your grade for this exam." Students will select a response ranging from 0-100 to indicate their predicted exam score. Exams for the course are multiple-choice and are scored using a Scantron device. The course instructor will provide me with students' actual exam scores on a 0-100% scale to measure calibration.

Schraw (2009) argues that absolute calibration, or the difference between predicted and actual exam scores, is the appropriate measure to use for intervention studies. Since SI can be thought of as an intervention, I will follow this standard by calculating the absolute difference between participants' predicted and actual exam scores. Thus, calibration scores may range from 0-100, with lower scores demonstrating greater calibration accuracy and a score of zero indicating perfect calibration.

I will use the absolute differences among predicted and actual scores in the SEM model. Then, students' bias scores will be used to examine the results descriptively. Bias scores are based on the direction of the calibration judgment with positive numbers reflecting overconfidence and negative numbers representing underconfidence (Hawthorne, Bol, & Pribesh, 2017). For example, if a student predicts he or she will earn an 80% but receives a 50%

on the exam, the overconfidence score would be +30. Conversely, a student who estimates he or she will produce an 80% but earns a 90% would have an underconfidence score of -10.

Final calibration. The measure of participants' final calibration is an endogenous variable that may be influenced by beginning calibration and SI attendance. At the end of the semester, students will be asked to respond to the same calibration question prior to their final exam: "On a scale of 0-100%, predict your grade for this exam." Students again will select a response ranging from 0-100 to indicate their predicted scores on the multiple-choice final exam, and the course instructor will provide me with students' actual exam scores on a 0-100% scale. Absolute calibration will be determined by calculating the differences in their predicted and actual exam scores.

Final calibration scores will be compared to beginning calibration scores to identify if students' calibration accuracy improved or decreased from the beginning to the end of the semester. For example, if Student A predicted he or she would earn an 80% but received a 50% on the first test, the beginning calibration score for that student would be 30. Then, if Student A predicted he or she would earn an 80% and received a 75% on the final exam, the student's absolute calibration score at the end of the semester would be 5. Since lower numbers indicate better calibration, Student A would have improved calibration accuracy from the beginning to the end of the term. As a comparison, if Student B also predicted he or she would earn an 80% but received a 50% on the first exam, this student also would have a beginning calibration score of 30. Then, if Student B predicted he or she would earn an 80% and received a 60% on the final exam, the final exam absolute calibration score would be 20. In this case, the final calibration score for Student B would be better than his or her beginning calibration score; however, the change would not be as significant as for Student A.

Self-Efficacy Scale

I will measure students' self-efficacy at the beginning and end of the semester when students are asked the exam calibration question. The pre- and post-tests will be used to measure beginning self-efficacy and final self-efficacy, respectively.

Beginning self-efficacy. Beginning self-efficacy is an exogenous variable that may influence SI attendance and final self-efficacy. I will use an existing scale from the Patterns of Adaptive Learning Scale (PALS) to measure participants' beginning self-efficacy (Midgley et al., 2000). Specifically, students will answer the five questions from the PALS Academic Efficacy scale with a minor adjustment of replacing "class" with "biology course" (See Appendix A). Each item asks students to rate themselves using a 5-point Likert scale ranging from 1 ("not at all true of me") to 5 ("very true of me"). The items ask students to reflect on their ability to (a) master skills taught, (b) figure out how to do the most difficult work, (c) do almost all the work by not giving up, (d) learn content even if it is hard, and (e) do even the hardest work by trying. A prior study found that the coefficient alpha for the Academic Efficacy scale is 0.78 (Midgley et al., 2000), and the construct validity for this scale has been supported by previous research that compared elementary and middle school students (Anderman & Midgley, 1997; Midgley, Anderman, & Hicks, 1995).

I will calculate the beginning self-efficacy variable for the path model by averaging each participant's responses to the five Likert-scale questions. Higher score averages will be indicative of higher self-efficacy.

Final self-efficacy. Final self-efficacy is an endogenous variable that may be predicted by beginning self-efficacy, SI attendance, and final calibration. Prior to the final exam, students again will be asked to respond to the five questions from the PALS Academic Efficacy scale.

Responses to the five questions will be averaged to produce the final self-efficacy variable. Final self-efficacy scores that are higher than beginning self-efficacy scores will reflect an increase in self-efficacy. For example, if a student has a beginning self-efficacy score of 1.8 and a final self-efficacy score of 3.9, his or her self-efficacy scores would signal a significant increase in self-efficacy towards the course content from the beginning to the end of the semester.

SI Attendance

SI attendance will be the total number of SI sessions attended by students each semester. SI leaders will collect student attendance electronically at the beginning and end of each session. The institution uses an online student data management system called Student Success Collaborative-Campus, which is managed by a company called the Education Advisory Board to capture student involvement in tutoring, SI, advising, and other related services (EAB Global, Inc., 2018). At the end of the semester, I will collect from the SI program an Excel report of student SI attendance, which will be matched with survey responses and demographic information using students' unique identification numbers (UINs).

Other Variables and Student Demographics

Several other path model variables and student demographics will be used in my study. I will request from the course instructor students' final course grades (See Appendix B for the request letter I will send to the course instructor). I also will request information from the institutional assessment office, including the path model variable of total SAT score and student demographic information for use as descriptive statistics, including gender, race/ethnicity, class standing, and major (Refer to Appendix C for the letter I will send to the institutional assessment office).

Final course grade. I will request from the course instructor students' final course grades in the general biology course on a 0-100% scale. This is an endogenous variable from the hypothesized path model that may be predicted by the total SAT score, SI attendance, final calibration, and final self-efficacy variables.

Total SAT score. I will request students' total SAT scores from the institution's assessment office. The scores will fit within a range from 400-1600. This is an exogenous variable in the path model that may predict SI attendance and final course grade.

Other student demographics. I will request other student demographic variables from the office of institutional assessment to serve as descriptive statistics of the study participants and the general biology class population. Specifically, demographic variables will include student gender, race/ethnicity, class standing (e.g., freshman), and major. This information will be presented in aggregate form.

Procedure

At the end of the second week of class, one week prior to the first exam, I will electronically distribute the student survey using a university-sponsored system called Qualtrics. Students will receive an email immediately prior to their class time during which I will introduce my study and ask students to complete the survey on their electronic devices. Students will be offered extra credit by the course instructor for completing the survey, and the instructor will offer an alternative extra credit assignment of completing problems from the back of the textbook to students who do not wish to complete the survey. Extra credit will be removed when calculating students' final course grades in the path analysis. In addition, students will be given the incentive of entering their name into a drawing for one of ten $10 Amazon gift cards, which I will award to randomly selected students at the end of the semester. Students will be able to enter their name into the drawing up to two times: once for the pretest and again for the posttest at the end of the semester.

The pretest survey will include a notification letter informing participants of the study's purpose, requirements, potential benefits and risks, voluntary nature, and assurance of confidentiality. The letter will notify them that, should they complete the assessment, I will match their responses to their demographic characteristics, grades, and SI attendance data. Participants also will be notified that the instructor and SI leaders will not have access to survey responses and that their responses will have no effect on their grades (Refer to Appendix D for the notification letter). Students will have the option to electronically consent to participate in the study prior to answering the survey questions. Students will be allotted time during class to complete the survey, and Qualtrics will be used to send them reminder emails each day, ending on the day of the exam.

Throughout the semester, SI leaders will host weekly sessions and ask students to sign-in to the session using an electronic kiosk. In the case of technical difficulties, SI leaders will collect student names and UINs via paper and enter information retroactively into the electronic system.

One week prior to the final exam, I will visit each class section to encourage students to complete the posttest survey emailed to them via Qualtrics immediately prior to their class period. Again, students will be informed of the purpose of the study and offered the incentives of extra credit and entering their name into a drawing for an Amazon gift card. The instructor will provide students with class time to complete the survey, and daily email reminders will be sent to students, ending on the date of the final exam.

Once final grades have been submitted, I will collect students' exam grades and final course grades from the instructor, requesting that extra credit for participation in the study be removed from the final grade calculations. In addition, SI attendance data will be collected from the electronic system, and additional student performance and demographic data will be requested from the institutional assessment office. Students' UINs will be used to merge all records, and IBM SPSS Statistics 24 and Mplus (v 7.3) will be used for all data analysis.

I will ensure confidentiality by asking students to use their UINs when completing both the pre- and post-assessments. In addition, participant information will be kept in a separate, password-protected database, and the data will be destroyed five years after the project is completed by deleting all associated files.

Data Analysis

This section outlines the analyses to be conducted once data have been collected. First, I describe how I will display the descriptive statistics. Then, I explain how I will conduct my path analysis.

Descriptive Statistics

I will begin my data analysis by examining the descriptive statistics of my collected data. The first set of data will be represented by a table depicting the frequencies and percentages of demographic factors, including gender, race, year, class standing, and major to assess the representativeness of the sample to the larger population of the general biology course. Another table will display the means, standard deviation, skew, and kurtosis for the path model's seven variables. Based on SI program data from previous semesters, it is unlikely that SI attendance will be normally distributed, so I will include another table detailing the frequencies and percentages for SI attendance data. Finally, a correlation matrix of the study variables will be provided.

Path Analysis

I will apply a path analysis with robust maximum likelihood estimation to answer my research questions using Mplus (v 7.3; Byrne, 2012). A maximum likelihood estimation will help account for the SI attendance data that likely will not be normally distributed (Byrne, 2012). Again, my hypothesized path model is in Figure 2. My cutoff value for statistically significant results will be $p < .05$, and only significant paths will be displayed in my results chapter.

I will use fit statistics recommended by Hu and Bentler (1999) to assess model fit, beginning with chi-square (x^2). The model will be considered a good fit if the chi-square statistics are small and insignificant; however, due to the sensitivity of x^2 to sample size and other issues, other indicators of model fit must be used. Assuming my sample size is over 250 participants, I will attempt to minimize errors by including the Tucker Lewis Index (TLI), root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR; Byrne, 2012; Hu & Bentler, 1999; Kline, 2016). The following cutoff values for these fit statistics are recommended: a TLI greater than .95, RMSEA less than .06, and SRMR of less than .05 (Byrne, 2012; Hu & Bentler, 1999).

While my hypothesized path model is based on theoretical and empirical literature, it is possible that my model will be rejected due to its poor fit with the sample data (Byrne, 2012). If this occurs, I will engage in a process known as "model generating" (Byrne, 2012, p. 8) by which I will release one path at a time and analyze the changes in chi-square to determine any statistically significant improvements in the model (Loehlin, 1998). I will describe in detail any such changes in my completed study.

Limitations

I anticipate limitations to my study. First, as with most human subject studies, self-selection bias is an issue for survey completion and SI session attendance. To control for selection bias, as well as other confounds, such as academic achievement, I will account for total SAT scores in my SEM model and use demographics to compare study participants with the entire class population.

Another threat to internal validity is social desirability, since the study uses self-report measures. I will control for this by administering the survey electronically to reduce students' fears that their course instructor, SI leader, or classmates may observe their responses. Confidentiality also will be assured to participants during in-class announcements and via the electronic notification letter.

Students will have the option to attend SI sessions led by two different SI leaders, which is another threat to internal validity. Fidelity will be enhanced by providing both SI leaders with an intensive pre-semester training and ongoing developmental opportunities, which have been recognized by the International Center for SI via the institution's SI program certification. In addition, part of the ongoing training of SI leaders involves session observations throughout the semester to ensure they are appropriately implementing the SI model. Finally, a strength of this study is that all students are taught by the same course instructor.

A final potential threat to internal validity is possible attrition of study participants. I will attempt to control for this by asking the instructor to offer students extra credit in the course and by allowing participants to enter their names into a gift card drawing for completion of the pre- and post-tests. A potential related problem could be low SI attendance, which could weaken the path model results. To combat this challenge, SI leaders will make periodic in-class announcements and send weekly reminders to students with session information. The course instructor also will be asked to encourage students to participate in SI.

The one-course, single-institution design of my study also threatens its external validity. I will account for this by providing institutional context and detailed demographic information for study participants and by cautioning readers on the generalizability of the study results to different contexts. Further studies also will be encouraged to duplicate the procedures of my research to build external validity over time.

REFERENCES

Anderman, E. M., & Midgley, C. (1997). Changes in achievement goal orientations and perceived classroom goal structures across the transition to middle level schools. *Contemporary Educational Psychology, 22*, 269–298.

Arendale, D. (1997). Supplemental Instruction (SI): Review of research concerning the effectiveness of SI from the University of Missouri-Kansas City and other institutions from across the United States. In S. Mioduski & G. Enright (Eds.), *Proceedings of the 17th and 18th Annual Institutes for Learning Assistance Professionals* (pp. 1-25). Tucson, AZ. Retrieved from ERIC Database. (ED457797).

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191-215.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: W. H. Freeman.

Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (1996). Multifaceted impact of self-efficacy beliefs on academic functioning. *Child Development, 67*, 1206-1222.

Bembenutty, H. (2007). Self-Regulation of learning and academic delay of gratification: Gender and ethnic differences among college students. *Journal of Advanced Academics, 18*(4), 586–616. http://doi.org/10.4219/jaa-2007-553

Bembenutty, H. (2009). Three essential components of college teaching: Achievement calibration, self-efficacy, and self-regulation. *College Student Journal, 43*(2), 562-570.

Blanc, R. A., DeBuhr, L. E., & Martin, D. C. (1983). Breaking the attrition cycle: The effects of Supplemental Instruction on undergraduate performance and attrition. *The Journal of Higher Education, 54*(1), 80-90.

Bol, L., & Hacker, D. J. (2001). A comparison of the effects of practice tests and traditional review on performance and calibration. *The Journal of Experimental Education, 69*, 133-151.

Bol, L., Hacker, D., Walck, C. & Nunnery, J. (2012). The effects of individual or group guidelines on the calibration accuracy and achievement of high school biology students. *Contemporary Educational Psychology, 37*, 280-287.

Bol, L., Riggs, R., Hacker, D., Dickerson, D., Nunnery, J. (2010). The calibration accuracy of middle school students in math classes. *Journal of Research in Education, 21*, 81-96.

Bowles, T. J., & Jones, J. (2003). An analysis of the effectiveness of Supplemental Instruction: The problem of selection bias and limited dependent variables. *Journal of College Student Retention, 5*(2), 235-243.

Bruno, P.A., Love Green, J. K., Illerbrun, S. L., Holness, D. A., Illerbrun, S. J. , Haus, K. A., … Sveinson, K. L. (2016). Students helping students: Evaluating a pilot program of peer teaching for an undergraduate course in human anatomy. *Anatomical Sciences, Education, 9*, 132-142.

Byrne, B. M. (2012). *Structural equation modeling with Mplus: Basic concepts, applications, and programming.* New York, NY: Routledge.

Chen, P. P. (2003). Exploring the accuracy and predictability of the self-efficacy beliefs of seventh-grade mathematics students. *Learning and Individual Differences, 14*, 79-92. doi: 10.1016/j.lindif.2003.08.003

Cohen, J. (1983). The cost of dichotomization. *Applied Psychological Measurement, 7*, 249-253.

Complete College America. (2014). *Guided pathways to success: Boosting college completion.* Retrieved from http://www.collegechangeseverything.org/events/scalingup-media/Guided_Pathways_to_Success_Summary-Complete_College_America.pdf

Curators of the University of Missouri. (2011). *The supervisor's guide to Supplemental Instruction.* Kansas City, MO.

Dancer, D., Morrison, K., & Tarr, G. (2015). Measuring the effects of peer learning on students' academic achievement in first-year business statistics. *Studies in Higher Education, 40*(10), 1808-1828. http://dx.doi.org/10.1080/03075079.2014.916671

DiGiacomo, G., & Chen, P. P. (2016). Enhancing self-regulatory skills through an intervention embedded in middle school mathematics curriculum. *Psychology in the Schools*, *53*, 601-616. https://doi-org.proxy.lib.odu.edu/10.1002/pits.21929

Dinsmore, D. L., & Parkinson, M. M. (2013). What are confidence judgments made of? Students' explanations for their confidence ratings and what that means for calibration. *Learning and Instruction, 24*, 4-14. https://doi.org/10.1016/j.learninstruc.2012.06.001

Douglas-Gabriel, D. (2015). Racial disparities in college major selection exacerbate earnings gap. *The Washington Post.* Retrieved from https://www.washingtonpost.com/news/grade-point/wp/2015/09/16/racial-disparities-in-college-major-selection-exacerbates-earnings-gap-3/?utm_term=.893b5377ef8d

EAB Global, Inc. (2018). Introducing higher education's first student success management system. Retrieved from https://www.eab.com/technology/student-success-collaborative

Elam, J. (2016). Academic support and mentoring: Supplemental Instruction overview. Retrieved from http://www.umkc.edu/asm/umkcsi/

Fayowski, V., & MacMillan, P. D. (2008). An evaluation of the Supplemental Instruction programme in a first year calculus course. *International Journal of Mathematical Education, 39*(7), 843-855. doi: 10.1080/00207390802054433

Fisher, J. E. (1997). *Effects of Supplemental Instruction on undergraduate academic achievement, motivational orientations, and learning strategies.* (Doctoral dissertation). Retrieved from ProQuest & Theses Dissertations Global. (304333981)

Flannelly, L. T. (2001). Using feedback to reduce students' judgment bias on test questions. *Journal of Nursing Education, 40*, 10-16.

Freire, P. (1993). Pedagogy of the oppressed. (M. B. Ramos, Trans.). New York: Continuum. (Originally published in English in 1972)

Garcia, E. (2006). *Supplemental Instruction, study habits, and the community college student.* (Master's thesis). Retrieved from ProQuest & Theses Dissertations Global. (304919971)

Gattis, K. W. (2002). Responding to self-selection bias in assessments of academic support programs: A motivational control study of Supplemental Instruction. *The Learning Assistance Review, 7*(2), 26-36.

Grier, T. (2004). Supplemental Instruction and noncognitive factors: Self-efficacy, outcome expectations, and effort regulation. *The Learning Assistance Review, 9*(2), 17-28.

Grimm, J., & Perez, T. (2017, April). *A longitudinal study of Supplemental Instruction's impact on anatomy & physiology course performance.* Poster session presented at the meeting of American Educational Research Association, San Antonio, TX.

Gutierrez, A., & Schraw, G. (2015). Effects of strategy training and incentives on students' performance, confidence, and calibration. *The Journal of Experimental Education, 83*, 386-404.

Hacker, D. J., Bol, L., Horgan, D., & Rakow, E. A. (2000). Test prediction and performance in a classroom context. *Journal of Educational Psychology*, 92, 160-170. doi:10.1037/0022-0663.92.1.160

Hacker, D. J., Bol, L., & Keener, M. C. (2008). Metacognition in education: A focus on calibration. In J. Dunlosky & R. A. Bjork (Eds.), *Handbook of metamemory and memory* (pp. 429-455). New York: Psychology Press.

Hacker, D. J., & Bol, L. (in press). Calibration and self-regulated learning: Making the connections. In J. Dunlosky & K. Rawson (Eds.). *Cambridge handbook on cognition and education.* Cambridge, UK: Cambridge University Press.

Hancock, G. R., & Mueller, R. O. (2010). *The reviewer's guide to quantitative methods in the social sciences*. New York, NY: Routledge.

Hawthorne, K., Bol, L., & Pribesh, S. (2017). Can providing rubrics for writing tasks improve developing writers' calibration accuracy? *Journal of Experimental Education, 85,* 689-708.

Hizer, S. E. (2010). *The Supplemental Instruction program: Student perceptions of the learning environment and impact on student academic achievement in college science at California State University, San Marcos.* (Doctoral dissertation). Retrieved from ProQuest & Theses Dissertations Global. (288307454)

Hong, J. C., Hwang, M. Y., Tai, K., H., & Chen, Y. L. (2014). Using calibration to enhance students' self-confidence in English vocabulary learning relevant to their judgment of over-confidence and predicted by smartphone self-efficacy and English learning anxiety. *Computers & Education, 72*, 313-322. http://dx.doi.org/10.1016/j.compedu.2013.11.011

Hopper, M. (2011). Student enrollment in a supplemental course for anatomy and physiology results in improved retention and success. *Journal of College Science Teaching*, *40*(3), 70-79.

Hu, L-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1-55.

Huff, J. D., & Nietfeld, J. L. (2009). Using strategy instruction and confidence judgments to improve metacognitive monitoring. *Metacognition and Learning, 4*, 161-176.

Hurley, M. A. (2000). *Video-based Supplemental Instruction (VSI): An interactive delivery system that facilitates student learning*. (Doctoral dissertation). Retrieved from ProQuest & Theses Dissertations Global. (9970741)

Hurley, M., Jacobs, G., & Gilbert, M. (2006). The basic SI model. In M. E. Stone, & G. Jacobs. (Eds.), *New Directions of Teaching and Learning, 2006*(106), 11-22. doi: 10.1002/tl.229

Jarvela, S., & Jarvenoja, H. (2011). Socially constructed self-regulated learning and motivation regulation in collaborative learning groups. *Teachers College Record, 113,* 350-374.

Johnson, D. W., Johnson, R. T., & Holubec, E. J. (1994). *The nuts and bolts of cooperative learning*. Edina, MN: Interaction Book.

Karabenick, S. A. (2003). Seeking help in large college classes: A person-centered approach. *Contemporary Educational Psychology*, *28*, 37-58. doi: 10.1016/S0361-476X(02)00012-7

Karabenick, S. A., & Berger, J. L. (2013). Help seeking as a self-regulated learning strategy. In H. Bembenutty & T. J. Cleary (Eds.), *Self-regulated learning across diverse disciplines: A tribute to Barry J. Zimmerman* (pp. 237-261). Information Age Publishing, Inc.

Karabenick, S. A., & Knapp, J. R. (1991). Relationship of academic help seeking to the use of learning strategies and other instrumental achievement behavior in college students. *Journal of Educational Psychology, 83*, 221–230.

Kitsantas, A., Winsler, A., & Huie, F. (2008). Self-regulation and ability predictors of academic success during college: A predictive validity study. *Journal of Advanced Academics, 20*(1), 42-68.

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). New York, NY: The Guilford Press.

Koenig, K., Schen, M., Edwards, M., & Bao, L. (2012). Addressing STEM retention through a scientific thought and methods course. *Journal of College Science Teaching, 41*(4), 23-29.

List, A., & Alexander, P. A. (2015). Examining response confidence in multiple text tasks. *Metacognition and Learning, 10*, 407-436. doi: 10.1007/s11409-015-9138-2

Loehlin, J. C. (1998). *Latent variable models: An introduction to factor, path, and structural analysis* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.

Mack, A. C. (2007). *Differences in academic performance and self-regulated learning based on level of student participation in Supplemental Instruction.* (Doctoral dissertation). Retrieved from ProQuest & Theses Dissertations Global. (304746773)

Martin, D. C., & Arendale, D. R. (1992). Review of research on Supplemental Instruction. In D. C. Martin & D. Arendale (Eds.), *Supplemental Instruction: Improving first-year student success in high-risk courses* (2nd ed.) (pp. 19-26). Columbia, SC: National Resource Center for the Freshman Year Experience and Students in Transition. Retrieved from ERIC Database (ED354839).

McGee, J. V. (2005). *Cognitive, demographic, and motivational factors as indicators of help seeking in Supplemental Instruction.* (Doctoral dissertation). Retrieved from ProQuest Dissertations & Theses Global. (305377445).

McGuire, S. Y. (2006). The impact of Supplemental Instruction on teaching students *how* to learn. In M. E. Stone, & G. Jacobs (Eds.), *New Directions of Teaching and Learning, 2006*(106), 3-10. doi: 10.1002/tl.228

Midgley, C., Anderman, E. M., & Hicks, L. (1995). Differences between elementary and middle school teachers and students: A goal theory approach. *Journal of Adolescence*, *15*, 90–113.

Midgley, C., Maehr, M. L., Hruda, L. Z., Anderman, E., Anderman, L., Freeman, K. E., … Urdan, T. (2000). *Manual for the Patterns of Adaptive Learning Scales.* Ann Arbor, MI: University of Michigan.

Miller, M. J., Lent, R. W., Lim, R. H., Hui, K., Martin, H. M., Jezzi, M. M., … Williams, K. (2015). Pursuing and adjusting to engineering majors: A qualitative analysis. *Journal of Career Assessment*, *23*(1), 48-63. doi: 10.1177/1069072714523084

Morrison, J., Bol, L., Ross, S., & Watson, G. (2015). Paraphrasing and prediction with self-explanation as generative strategies for learning science principles. *Education Technology: Research and Development, 63,* 861-882. doi: 10.1007/s11423-015-9397-2

Nasr, P. (2012). A report on case study application in an undergraduate anatomy and physiology course. *AURCO Journal*, *18*, 123-139.

National Science Foundation. (2011). *Empowering the nation through discovery and innovation (Fiscal Year 2011-2016).* Retrieved from www.nsf.gov/news/strategicplan/nsfstrategicplan_2011_2016.pdf

Newman, R. S. (2008). The motivational role of adaptive help seeking in self-regulated learning. In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 315-337). New York, NY: Lawrence Erlbaum Associates.

Nietfeld, J. L., Cao, L., & Osborne, J. W. (2006). The effect of distributed monitoring exercises and feedback on performance, monitoring accuracy, and self-efficacy. *Metacognition and Learning, 1*, 159-179. doi:10.1007/s10409-006-9595-6.

Ning, H. K., & Downing, K. (2010). The impact of Supplemental Instruction on learning competence and academic performance. *Studies in Higher Education, 35*(8), 921-939. doi: 10.1080/03075070903390786

Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research, 66*(4), 543-578.

Pajares, F. (2002). Gender and perceived self-efficacy in self-regulated learning. *Theory into Practice, 41*(2), 116-125.

Perez, T., Cromley, J. G., & Kalan, A. (2014). The role of identity development, values, and costs in college STEM retention. *Journal of Educational Psychology, 106*(1), 315-329. doi: 10.1037/a0034027

Peverly, S. J., Brobst, K. E., Graham, M., & Shaw, R. (2003). College adults are not good at self-regulation: A study on the relationship of self-regulation, note taking, and test taking. *Journal of Educational Psychology, 95*, 335-346.

Pew Research Center (2014). *Center on budget and policy priorities: Racial, ethnic wealth gaps have grown since great recession.* Retrieved from http://www.pewresearch.org/fact-tank/2014/12/12/racial-wealth-gaps-great-recession/

Rabitoy, E. R., Hoffman, J. L., & Person, D. R. (2015). Supplemental Instruction: The effect of demographic and academic preparation variables on community college student academic achievement in STEM-related fields. *Journal of Hispanic Higher Education, 14*(3), 240–255. doi: 10.1177/1538192714568808

Reid, A. J., Morrison, G. R., & Bol, L. (2016). Knowing what you know: Improving metacomprehension and calibration accuracy in digital text. *Educational Technology Research and Development, 65*(1), 29-45. doi: 10.10007/s11423-016-9454-5.

Schraw, G. (2009). Measuring metacognitive judgments. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 415-429). New York, NY: Routledge.

Schudde, L. T., & Goldrick-Rab, S. (2016). Extending opportunity, perpetuating privilege: institutional stratification amid educational expansion. In M. N. Bastedo, P. G. Altbach, & P. J. Gumport (Eds.), *American higher education in the 21st century* (pp. 345-374). Baltimore, MD: Johns Hopkins University Press.

Schunk, D. H. (1990). Goal setting and self-efficacy during self-regulated learning. *Educational Psychologist, 25*(1), 71-86.

Schunk, D. H. (1991). Self-efficacy and academic motivation. *Educational Psychologist, 26*(3), 207-231.

Schunk, D. H. (2012). Social cognitive theory. In K. R. Harris, S. Graham, & T. Urdan (Eds.), *APA educational psychology handbook, Vol. 1: Theories, constructs, and critical issues.* (pp. 101-123). American Psychological Association. doi: 10.1037/13273-005

Schunk, D. H., & Pajares, F. (2005). Competence perceptions and academic functioning. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 85-104). New York, NY: The Guilford Press.

Shaughnessy, J. J. (1979). Confidence-judgment accuracy as a predictor of test performance. *Journal of Research in Personality, 13*, 505-514.

Sitzmann, T., & Ely, K. (2011). A meta-analysis of self-regulated learning: What we know and where we need to go. *Psychological Bulletin, 137*(3), 441-442. doi: 10.1037/a0022777

Smith, D. G. (2016). The diversity imperative: Moving to the next generation. In M. N. Bastedo, P. G. Altbach, & P. J. Gumport (Eds.), *American higher education in the 21st century* (pp. 375-400). Baltimore, MD: Johns Hopkins University Press.

StataCorp LLC (2018). *SEM intro 4: Substantive concepts.* Retrieved from https://www.stata.com/manuals13/semintro4.pdf

Terrion, J.L., & Daoust, J.L. (2012). Assessing the impact of supplemental instruction on the retention of undergraduate students after controlling for motivation. *Journal of College Student Retention, 13*(3): 311-327. doi: 10.2190/CS.13.3.c

UMKC SI (2018). *The International Center for Supplemental Instruction.* Retrieved from http://info.umkc.edu/si/

Usher, E. L. (2009). Sources of middle school students' self-efficacy in mathematics: A qualitative investigation. *American Educational Research Journal, 46*(1), 275–314. doi: 10.3102/0002831208324517

Usher, E. L. (2016). Personal capability beliefs. In L. Corno & E. M. Anderson (Eds.), *Handbook of Educational Psychology* (pp. 146–159). New York, NY: Routledge.

Visor, J. N., Johnson, J. J., & Cole, L. N. (1992). The relationship of Supplemental Instruction to affect. *Journal of Developmental Education*, *16*(2), 12-14, 16-18.

Watters, J. J., & Ginns, I. S. (1997, March). *Peer assisted learning: Impact on self-efficacy and achievement*. Paper presented at the meeting of American Educational Research Association Conference, Chicago, IL.

Widmar, G.E. (1994). Supplemental Instruction: From small beginnings to a national program. In D. C. Martin and D. Arendale (Eds.), *Supplemental Instruction: Increasing Achievement and Retention* (pp. 3-10). San Francisco, CA: Jossey Bass, Inc.

Zaritsky, J. S., Toce, A. (2006). Supplemental Instruction at a community college: The four pillars. In M. E. Stone & G. Jacobs (Eds.), *New Directions of Teaching and Learning, 2006*(106), 11-22. doi: 10.1002/tl.229

Zerger, S. (2008). Theoretical frameworks that inform the Supplemental Instruction model. In M. E. Stone & G. Jacobs (Eds.), *Supplemental Instruction: Improving First-Year Student Success in High-Risk Courses* (pp. 21-28). Columbia, SC: National Resource Center for the Freshman Year Experience and Students in Transition.

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts & P. R. Pintrich (Eds.), *Handbook of Self-Regulation* (pp. 13-39). New York: Academic Press.

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice, 41 (2*), 64-70.

Zimmerman, B. J., Bandura, A., & Martinez-Pons, M. (1992). Self-motivation for academic attainment: The role of self-efficacy beliefs and personal goal setting. *American Educational Research Journal, 29*(3), 663-676.

APPENDIX A

PALS Academic Efficacy Scale

1. I am certain I can master the skills taught in this biology course.

1	2	3	4	5
NOT AT ALL TRUE		SOMEWHAT TRUE		VERY TRUE

2. I'm certain I figure out how to do the most difficult coursework in this biology course.

1	2	3	4	5
NOT AT ALL TRUE		SOMEWHAT TRUE		VERY TRUE

3. I can do almost all of the work in this biology course if I don't give up.

1	2	3	4	5
NOT AT ALL TRUE		SOMEWHAT TRUE		VERY TRUE

4. Even if the work in this biology course is hard, I can learn it.

1	2	3	4	5
NOT AT ALL TRUE		SOMEWHAT TRUE		VERY TRUE

5. I can do even the hardest work in this biology course if I try.

1	2	3	4	5
NOT AT ALL TRUE		SOMEWHAT TRUE		VERY TRUE

APPENDIX B

Course Instructor Data Request Letter

Dear Dr. Mills:

My name is Jenn Grimm, and I have worked at ODU as the Director of the Peer Educator Program since September 2015. In addition, I am currently a Ph.D. student in the Higher Education program at ODU. I am requesting your assistance with my research study, which will examine the effects of students' participation in Peer-Assisted Study Sessions (PASS) on self-efficacy and calibration accuracy. My dissertation is titled *Supplemental Instruction, Calibration, and Self-Efficacy: A Path Model Analysis*.

I would like to invite students in your BIOL 121N course to participate in my study during the fall 2018 semester. Specifically, I am reaching out to you to request the following opportunities:

1. **To distribute to your students an electronic survey through Qualtrics:** This survey will be distributed one week prior to the first and final exams. I request that you allow me 5-10 minutes of your class times during these days to introduce the study to your students and to have them complete the brief survey.
2. **To offer extra credit to your students who complete each survey:** The extra credit will be offered to students at two separate times, once for the pretest and again for the posttest. Students should be given the option of completing an alternative assignment to receive extra credit, should they choose to not participate in the study.
3. **To provide me with access to students' final course grades and exam scores:** I will need access to the final course grades and students' performance on the first and final exams on a 0-100% scale. The final course grade calculations will need to have the extra credit points for study participation removed from students' scores.

Would you be willing to grant me the above opportunities to assist me with my dissertation research? I will be happy to share my dissertation proposal with you and answer any questions you may have. Thank you in advance for your time and support.

Sincerely,
Jenn Grimm

APPENDIX C

Office of Institutional Effectiveness & Assessment Data Request Letter

Dear Dr. Parades:

My name is Jenn Grimm, and I have worked at ODU as the Director of the Peer Educator Program since September 2015. In addition, I am currently a Ph.D. student in the Higher Education program at ODU. I am requesting your assistance in my research study, which will examine the effects of students' participation in Peer-Assisted Study Sessions (PASS) on self-efficacy and calibration accuracy. My dissertation is titled *Supplemental Instruction, Calibration, and Self-Efficacy: A Path Model Analysis*.

I am writing to request performance and demographic information for students enrolled in BIOL 121N during the fall 2018 semester. Specifically, I am reaching out to you to request the following information for these students:

1. Total SAT scores
2. Gender
3. Race/ethnicity
4. Class standing
5. Major

Would you be willing to provide me the above information to assist me with my dissertation research? I will be happy to share my dissertation proposal with you and answer any questions you may have. Thank you in advance for your time and support.

Sincerely,
Jenn Grimm

APPENDIX D

Student Notification Letter

Dear Student:

I am a doctoral student at Old Dominion University. My study focuses on how your learning behaviors may influence your decision to attend PASS (Peer-Assisted Study Sessions) and how PASS may influence your learning behaviors. I need your help to improve student learning support opportunities. This brief survey should only take you two minutes to complete.

If you decide to complete this survey, you can receive extra credit from Dr. Mills. You may also enter your name into a drawing for one of ten $10 Amazon gift cards.

There are no known risks associated with this study. The researchers will maintain strict confidentiality. You will not be asked to provide your name but instead to use your unique identification number (UIN). Upon completing this survey, your UIN will be used to match your responses with your PASS attendance and information from your student records. The results of this study may be used in reports, presentations, and publications, but information will be presented in aggregate form and you will not be identified.

Your participation is voluntary. You can decline to complete the survey. Your responses will not be shared with the course instructor or SI leaders. There is no way your participation or responses will affect your grade or have any other consequences for you, so we do hope you decide to help us!

If you have any questions about this study, please contact Jenn Grimm at jgrimm@odu.edu, Dr. Chris Glass (Dissertation Committee Chair) at crglass@odu.edu, or Dr. Jill Stefaniak (Chair of the Human Subjects Review Committee for the Darden College of Education) at jstefani@odu.edu. Thank you very much for your consideration.

Sincerely,
Jenn Grimm