Summer 2019

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

Jennifer Leigh Grimm
Old Dominion University, jennhgrimm@gmail.com

SUPPLEMENTAL INSTRUCTION, CALIBRATION, AND SELF-EFFICACY: A PATH

MODEL ANALYSIS

by

Jennifer Leigh Grimm
B.B.A. May 2009, Ohio University
M.Ed. May 2011, Ohio University

A Dissertation Submitted to the Faculty of
Old Dominion University in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

EDUCATION – HIGHER EDUCATION CONCENTRATION

OLD DOMINION UNIVERSITY
August 2019

Approved by:

Christopher R. Glass (Director)

Tony Perez (Member)

Linda Bol (Member)

ABSTRACT

SUPPLEMENTAL INSTRUCTION, CALIBRATION, AND SELF-EFFICACY: A PATH MODEL ANALYSIS

Jennifer Leigh Grimm
Old Dominion University, 2019
Director: Dr. Christopher R. Glass

Many students preparing for careers in the fields of science, technology, engineering, and mathematics (STEM) are unable to persist past entry-level courses to complete their college degrees. As a result, many higher education institutions have implemented intervention programs, like Supplemental Instruction (SI), to help students master course content and gain the self-regulated learning (SRL) behaviors necessary for success in challenging STEM courses. Numerous studies have demonstrated that SI attendance is correlated with improved course grades; however, few studies have examined the effect of SI attendance on students' SRL behaviors, like self-efficacy and calibration, which may explain students' academic achievement throughout college.

The present study examined if students' pre-existing self-efficacy beliefs and calibration accuracy predicted their decisions to attend SI. In addition, the study explored if SI attendance had a direct effect on students' final self-efficacy, calibration, and course grades. Students in a fall semester general biology course for science majors were invited to participate in the study, and 320 students completed the pre- and post-test survey. The surveys measured beginning and final self-efficacy using the Academic Efficacy Scale from the Patterns of Adaptive Learning Scale, and calibration was measured by asking them to predict their first and final exam scores.

A path model was analyzed in Mplus via robust maximum likelihood estimations using pre- and post-test results and students' total SAT scores, SI attendance, and final course grades.

The results indicated that participants with lower self-efficacy were more likely to attend SI; however, students' beginning calibration accuracy did not predict their SI attendance. Findings also indicated that SI attendance did not predict final self-efficacy or calibration accuracy, but attending SI had a modest, direct effect on participants' final course grades. Final self-efficacy and calibration accuracy also predicted final course grades.

The results of this study demonstrate a need to explore additional SRL variables that may be influenced by SI. In addition, the present study validates the value of SI as an academic support program to raise course grades. Finally, potential course-level instructional strategies are offered for improving students' self-efficacy and calibration accuracy to support STEM degree persistence.

ACKNOWLEDGEMENTS

First and foremost, I want to thank my husband, Dr. Kevin Grimm, who has supported me every step of the way throughout my Ph.D. program journey and my career. I am also thankful to our sweet son, David, whose arrival into this world gave me the perspective I needed about what really matters in life. I need to thank my parents, Richard and Rhonda Haviland, and my big brother Matt, who have always believed in my potential and told me I could do anything I set my mind to do. My parents-in-law, Ron and Nancy Grimm, have also been incredibly supportive, especially in the time they have spent with our son David on those evenings and weekends when his mom had to stay late to work on schoolwork.

In addition, I could not have asked for a better dissertation committee. Dr. Chris Glass is the best chair, advisor, professor, and graduate program director, and Old Dominion University (ODU) is immensely fortunate to have him. I want to thank Chris for taking me under his wing when I started the Higher Education Ph.D. program at ODU as a transfer student in fall 2015. He helped me elevate my dissertation to new heights, and he was incredibly supportive every step of the way. My other committee members, Dr. Tony Perez and Dr. Linda Bol, taught me so much about educational psychology and research methods in their courses and offered additional, helpful guidance and feedback throughout the dissertation process. I also want to extend a sincere thank you to the anonymous Biology instructor who helped with my research.

I also need to thank my classmates and colleagues at ODU who cheered me along every step of the way. There are too many wonderful people to name, but please know that, whether you were a classmate or colleague who took an interest in my education, I appreciate you so much! Thank you also to my Executive Director, Lisa Mayes, who provided me with the professional space to pursue my Ph.D. while working full-time and prioritizing my family.

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; Rask, 2010), 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 relates 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; Rask, 2010; 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). As a result, 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 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 from 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 then use content learned in class and via

course assignments to plan weekly, optional, out-of-class group study sessions to provide

students with additional opportunities to review 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 an 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. If students develop improved SRL behaviors through SI, institutions may be more willing to invest in SI, which 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 improvements in self-efficacy and calibration may have on students' final course grades.

Purpose Statement

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

Research Questions

Three research questions guided the 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?

3. What is the indirect effect of SI attendance on final course grades through calibration and

    self-efficacy?

Overview of Methodology

I employed a non-experimental correlational design and used path modeling to answer

the research questions. The exogenous (or independent) variables included in the hypothesized

path model were total SAT score, beginning calibration, and beginning self-efficacy. The

endogenous (or dependent) variables were SI attendance, final calibration, final self-efficacy, and

final course grade. I recruited from approximately 540 potential participants from an

introductory undergraduate 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 were administered to participants prior to the first and

final course exams. SI attendance was collected from the SI program. The course instructor

provided final course grades and exam grades, and the institutional assessment office shared total

SAT scores and student demographic variables.

I applied a path analysis with robust maximum likelihood estimation to answer my

research questions using Mplus (v 7.3; Byrne, 2012). Fit criteria recommended by Hu and

Bentler (1999) were used to assess model fit, including chi-square (X^2), Comparative Fit Index

(CFI), root mean square error of approximation (RMSEA), and standardized root mean square

residual (SRMR). In addition, I checked my data to make sure it met the assumptions for

multivariate procedures. While my hypothesized path model was based on theoretical and

empirical literature, my original model was rejected due to its poor fit with the sample data (Byrne, 2012). I engaged in a process known as "model generating" (Byrne, 2012, p. 8) by which I utilized modification indices to identify and determine statistically significant improvements to develop an adjusted path model (Loehlin, 1998). My final path model was generated to display significant paths among the model variables.

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, 2019) and involves measuring how a person's perception of their performance matches his or her actual performance (Hacker et al., 2008).

Delimitations

I selected several delimitations to guide the scope of my study. First, I chose 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 examined a Supplemental Instruction program at a four-year research institution because it was an accessible sample and STEM education is important at the institution. I also 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. In addition, I chose to use Zimmerman's theory of SRL due to its use in other research studies that have examined SI and SRL. To streamline the SEM model further, I chose to use total SAT score as a predictor of prior achievement; however, other indices of achievement, including high school GPA, could have been used. I also selected to use final course grade, instead of final exam grade, as an endogenous variable due to its common use in SI research. Finally, I further chose to limit my study by not including in my path model demographic characteristics such as gender or race/ethnicity. I chose many of these delimitations to limit the number of variables used within the path model to reduce the number of required participants and to increase the likelihood of achieving statistically significant relationships among the variables.

Significance of the Study

This quantitative study contributes to SI program and educational psychology research in several ways. First, my research adds to and addresses 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 adds to the limited empirical literature that has examined how self-efficacy and calibration interact with and influence one another. Finally, this affords 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 related to 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 examine 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 describe 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 undergraduate classes in response to first- and second-year student attrition rates of 40 percent.

The SI model is 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 race/ethnicity and prior academic achievement. Second, SI participants succeed at higher rates than non-participants do, regardless of race/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.

The SI literature outlines 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 withdraw 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 than non-SI participants (Martin & Arendale, 1992).

SI impact on grades and DFW rates. Many SI studies have found 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 multiple regression (Grimm & Perez, 2017; Rabitoy et al., 2015).

Most studies have found these positive results, even though SI participants had significantly lower SAT (Peterfreund, Rath, Xenos, & Bayliss, 2008), ACT (Hensen & Shelley, 2003), and AAR (ACT Aptitude Rating; Moore & LeDee, 2006) scores than non-SI participants. The one exception was a study by Congos and Mack (2005) in which there were no significant differences in SAT scores between students in the SI and non-SI groups.

In addition, some researchers have looked for potential differences in the effects of SI attendance on students' academic performance based on gender (Fayowski & MacMillan, 2008; Mack, 2007) and race/ethnicity (Mack, 2007). In these studies, the researchers found no

statistically significant differences in the effects of SI attendance on academic performance based on gender (Fayowski & MacMillan, 2008; Mack, 2007) or race/ethnicity (Mack, 2007).

While most studies examined a single institution, a national SI field study was conducted from 1982-1996 on 270 institutions supporting over 505,000 students in nearly 5,000 courses (Arendale, 1997; Martin & Arendale, 1992). Aggregating this institutional data, the average final course grade in SI-supported courses for SI participants was 2.42 compared with an average course grade of only 2.09 for non-SI participants. Similarly, the DFW rates for SI participants was only 23.1 percent versus 37.1 percent for non-SI participants. These results were statistically significant (Arendale, 1997; Martin & Arendale, 1992).

In addition, two studies provided a breakdown of the UMKC SI program's impact on course grade by examining differences between SI and non-SI participants across top and bottom student quartiles determined by institutional admissions standards (Arendale, 1997; Blanc et al., 1983). Blanc et al. (1983) found statistically different final course grade averages between SI and non-SI participants across top and bottom quartiles at UMKC in Spring 1980. Students in the top quartile (n=149) who attended SI had a 3.10 average final course grade compared to a 2.30 average among non-SI participants. The average final course grades among SI and non-SI participants in the bottom quartile (n=75) was 1.72 and 0.88, respectively. Arendale (1997) also shared statistically significant data from a study conducted in 1989-1990 with 1,628 student participants. Students in the top quartile who participated in SI had an average final course grade of 3.29 compared with a 2.83 average for non-SI participants. Similarly, students in the bottom quartile who participated in SI had higher final course grade averages than non-SI participants (2.10 vs. 1.77). 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 statistically significant 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 who did not participate in SI. 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 examined 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 in which they provided enough descriptive detail about the SI programs being examined to indicate the programs followed the SI model.

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 researcher defines the SI treatment group differently in each study. For example, some researchers have placed 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) to be included in the SI participants' group. Thus, there is a great deal of variability in how researchers 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 or even nine 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 with

students at a minority-serving institution. 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 (A&P) courses. Results indicated that SI attendance in both courses had a significant positive effect on course grades, even after controlling for prior achievement. In addition, there were indirect effects of attending SI on course grades. Specifically, they found that students who attended more SI sessions in the first semester course (A&P I) were more likely to attend more SI sessions in their second semester (A&P II), leading to higher achievement in A&P II. The researchers also discovered that the indirect effects of students achieving higher grades in A&P I because of attending more SI sessions in A&P I led to higher course grades in A&P II. More studies 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. Through a thorough examination of the literature, I identified 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, et al., 1992; Watters & Ginns, 1997). I will review these studies later in this literature review. 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

A commonly used model for describing SRL processes is Zimmerman's (2000, 2002) three-phase model, which is derived 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 phases: forethought, performance, and self-reflection. Figure 1 presents a visual representation of Zimmerman's (2002) SRL model.

Performance Phase

Self-Control
Imagery
Self-instruction
Attention focusing
Task Strategies

Self-Observation
Self-recording
Self-experimentation

Forethought Phase

Task Analyis
Goal-setting
Strategic planning

Self-Motivation Beliefs
Self-efficacy
Outcome expectations
Intrinsic interest/value
Learning goal orientation

Self-Reflection Phase

Self-Judgment
Self-evaluation
Causal attribution

Self-Reaction
Self-satisfaction/affect
Adaptive/defensive

Figure 1. From: 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 present 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 successes or 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) investigated 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. This was likely due to the small sample size. Generalizability of this study is limited further 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, and 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 employed a quasi-experimental study in which

students in existing courses were assigned to mandatory SI treatment and control groups 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. The author opted to make SI sessions mandatory for certain course sections to

control for self-selection bias. 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 (the motivation

scale combines into one construct intrinsic/extrinsic goal orientation, task value, control of

learning beliefs, self-efficacy, and test anxiety). 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; however, 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 a 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 positive between student participation in SI on 7 of the 11 measured variables for the high-engagement group, 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 studies investigating both SI and SRL. 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 I will later argue is present in all three

phases of SRL (Hacker & Bol, 2019). Next, I discuss the theoretical and practical implications

for examining self-efficacy and calibration in my research study, including why I chose these

specific SRL constructs.

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 accuracy;

Bandura, Barbaranelli, Caprara, & Pastorelli, 1996; Schunk, 2012).

In addition, self-efficacy research has provided several implications for classroom

instructors. First, an emphasis on building students' mastery experiences is essential, since

performance-based information has the strongest influence on students' self-efficacy (Bandura,

1977; Schunk, 1991). Pajares (2002) suggests that teachers can do this by providing students

with tasks that are both challenging and meaningful but that they are capable of mastering. It is

paramount that teachers provide support and encouragement to students as they work on these

tasks but provide enough autonomy for students to engage independently in accomplishing these

tasks. Schunk (1991) also recommends that faculty enhance students' self-efficacy by providing

feedback for early successes, especially when students have had to put forth effort to accomplish their tasks, and rewards may also be used in these efforts. Another simple practice is for instructors to point out explicitly to students what they have learned already in their course and how the next topic will draw on their prior knowledge (Schunk, 2012). Modeling for students is also critical, including showing students the value of specific SRL and learning strategies, as well as demonstrating that it is okay to make mistakes (Pajares, 2002; Schunk, 1991). Finally, instructors can build-up students' self-efficacy by helping them set appropriate learning goals. Specifically, students should be encouraged to set short-term, proximal goals that include specific performance standards and start off easy before becoming progressively more difficult (Pajares, 2002; Schunk, 1991). 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). Leaders also are trained to encourage students to participate in activities in a safe, low-risk environment, thus allowing positive social persuasions to take place (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 from the beginning to the end of the term. 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), Study Behaviors Inventory (SBI; Garcia, 2006), Motivated Strategies for Learning Questionnaire (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 already 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 major 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 findings.

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 affect positively both calibration accuracy and academic performance (Hacker & Bol, 2019).

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' calibration predictions. 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 are, 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 (2019) 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 initiated 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 performances, 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 undergraduate students 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

noteworthy 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 the program.

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. 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 seek the help, 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. If Karabenick and Knapp's findings carry over to students' help-seeking behaviors in SI, the results of my study should indicate that students with poor calibration scores and low self-efficacy will be less likely to attend SI than their peers with accurate calibration and high self-efficacy.

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). 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).

The above 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 that I already outlined reveals also 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, 2019). Thus, an important component of SI that makes it an effective academic intervention strategy is that it supports students during all stages of the help-seeking and SRL processes. In addition, as I outlined in the previous section, SI 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, as I have argued previously, the activities engaged in during SI sessions should allow students to calibrate more accurately their knowledge and skills with their subsequent academic performance. This is particularly characteristic of the beginning SI session activities that help students identify what they already do or do not know, as well as the closing session activities that involve self-reflective practices. This 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 learning the course material and influence their decisions to continue to participate in SI sessions if they need the additional help (or not attend SI if they do not require

the help). In summary, students with proficient calibration accuracy should be the students most

likely to engage in appropriate and adaptive help-seeking behaviors through SI session

attendance.

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 also are 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' goal orientations, which can be performance-based (i.e., doing

better than other students to appear "smarter") or mastery-focused (i.e., being interested in

learning for understanding; see Table 2 for a simplified version of Newman's model).

Table 2

Help Seeking & Self-Efficacy

Is Help Necessary?	Action	
	Seek Help	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 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. If Newman's model were to be applicable to help-seeking and SI, I would anticipate that participants in my study who are high-achieving students would be more likely to attend SI if they had low self-efficacy and poor calibration accuracy (i.e., they would attend SI even though they may not need the extra help). I also would anticipate that low-achieving study participants would be more likely to attend SI if they have high self-efficacy and

strong calibration accuracy (i.e., they would attend SI because they have recognized that they will benefit from the help).

<div align="center">**Justification for Study**</div>

When considering a well-established higher education academic support program like Supplemental Instruction, there is value in knowing that there are links among 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 also is 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 of 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. While some studies

have examined SI and self-efficacy, there are several limitations to these studies. First,

researchers have inconsistently defined the SI treatment group. Most studies also have included

instruments that are too global in nature to measure 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 did 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 predicting 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 three 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?

3. What is the indirect effect of SI attendance on final course grades through calibration and self-efficacy?

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 used to address my research questions and hypotheses, including the study design, participants and context, measures, procedure, and data analysis.

Research Questions

Again, the following research questions guided 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?

3. What is the indirect effect of SI attendance on final course grades through calibration and self-efficacy?

Hypotheses

The first research question addressed 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 indicated that students with low self-efficacy were more likely to participate in SI (Hizer, 2010; McGee, 2005). However, in the help-seeking literature, Newman (2008) suggested from a theoretical perspective that students with high self-efficacy and the ability to predict their need for help would participate in an academic support

intervention, like SI, if they determined it was needed (or they would not participate if they did not determine that it was needed). Thus, prior to conducting this study, it was unknown if self-efficacy would be correlated positively or negatively with SI attendance. In addition, since no existing research had looked at calibration and help-seeking behaviors, it was unknown if calibration accuracy would predict students' SI attendance.

The second research question examined 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, it was anticipated that SI attendance would predict a positive increase in final calibration accuracy; however, it was unclear if SI attendance would predict a positive increase in 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 were useful for this hypothesis. Hacker and Bol (2019) 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 aligns with the three phases of SRL, I expected 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 seemed likely that those who participated in SI would perform better and have better final calibration accuracy than those who did not participate.

The second research question also examined how SI attendance may predict final self-efficacy. The effect of SI attendance on final self-efficacy was unclear prior to collecting data. It seemed likely that all students in the course would have lower self-efficacy by the end of the

semester, but the decrease in self-efficacy would be less dramatic for frequent SI participants (Hizer, 2010). However, it also was possible that the effect of SI attendance on final self-efficacy would not be detectable by the end of the semester (Fisher, 1997; Garcia, 2006; Grier, 2004; Watters & Ginns, 1997). Final calibration also may have affected 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.

Lastly, research question two asked about the direct effects SI attendance could have on final course grade. Previous SI research indicated that SI attendance would predict an increase in students' final course grades (e.g., Grimm & Perez, 2017; Rabitoy et al., 2015).

The third research question asked whether changes in calibration and self-efficacy because of SI attendance would indirectly influence final course grades. Nietfeld et al. (2006) suggest that even modest metacognitive monitoring interventions, like SI, can improve students' calibration accuracy, self-efficacy, and academic performance. Therefore, I predicted that increases in calibration and self-efficacy would have indirect positive effects of SI attendance on final course grade.

Research Design and Path Model

I employed 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 typically to represent causal processes among multiple variables (Byrne, 2012). It is important to note that this study was non-experimental in design (i.e., there was no random assignment of students to the SI treatment and non-SI treatments); therefore, study results indicated correlational relationships rather than causation.

The SEM 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 an inference method that uses three inputs to generate three outputs. The three inputs, which were present in the current study, included: (1) qualitative causal (or in this case, correlational) hypotheses based on theory or empirical findings, (2) questions about causal (or correlational) relations among study variables, and (3) data that are often used from non-experimental designs. The three outputs generated in the SEM included: (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 was 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, SEM allowed for the identification of potential direct and indirect effects of SI attendance on students' final course grade. Specifically, I conducted a path model analysis, which, according to Kline (2016), is a commonly used model in SEM. A path model was useful for the present study because each variable could be described with a single measure (e.g., beginning self-efficacy), and it was anticipated that the sample size may not have been large enough to warrant including a measurement model (Kline, 2016). Figure 2 depicts my hypothesized path model.
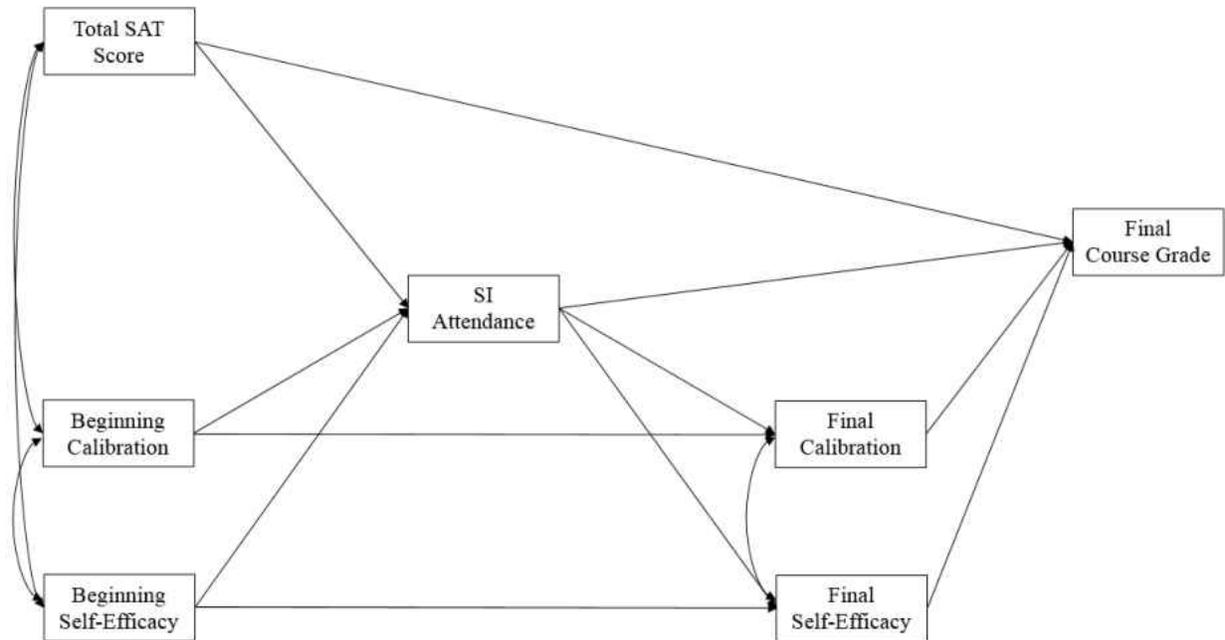
Figure 2. Hypothesized path model tested to determine relationships among total SAT score, beginning calibration and self-efficacy, SI attendance, final calibration and self-efficacy, and final course grade.

Participants

The SI program at the institution used for this study supports students in an introductory biology course for science majors each fall semester. One instructor teaches three sections of this course, and 529 students were enrolled in the course, across all three sections, at the beginning of fall 2018. Among these students, 422 (80%) participated in the pretest survey and completed the first exam. There were 47 students who withdrew from the course, resulting in 482 students at the close of the semester. Of the 482 students enrolled at the end of the semester, 320 students completed both the pretest and posttest surveys and first and final exams for a 66% class participation rate among students enrolled at the end of the semester. Table 3 provides an overview of the study participants, including gender, race/ethnicity, and class standing. Most of the study participants were female (71.9%), African-American or Caucasian (35.9% and 35.0%, respectively), and freshmen (60.9%).

Table 3

Characteristics of Study Participants

Characteristic	Subcategory	n	p
Gender			
	Female	230	71.9
	Male	90	28.1
Race/Ethnicity			
	African-American	115	35.9
	Caucasian	112	35.0
	Two or more races	31	9.7
	Hispanic/Latino	28	8.8
	Asian	21	6.6
	Native Hawaiian/Pacific Islander	0	0.0
	Non-resident aliens	2	0.6
	Unknown	11	3.4
Class Standing			
	Freshman	195	60.9
	Sophomore	76	23.8
	Junior	36	11.3
	Senior	10	3.1
	Graduate/Unclassified	3	0.9

Hancock and Mueller (2010) recommend having a minimum of five participants per parameter in a path model to obtain trustworthy maximum likelihood (ML) estimation, while Kline (2016) recommends at least a 10:1 sample-size-to-parameters ratio. The number of parameters in a hypothesized model 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) has up to 28 parameters (7*8/2), which means that my study ideally should have achieved a minimum of 140-280 participants (Hancock & Mueller, 2010; Kline, 2016). Since 320 students participated in the study, I exceeded this goal.

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 56.3% female and 43.7% 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. I chose to conduct my research at this institution because its diverse student population mirrors the demographics of many other diverse US institutions and because it has achieved SI program certification recognition by the International Center for SI at the University of Missouri-Kansas City. I also selected this program because I had convenient access to much of the needed data.

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 supported three sections of the general biology course that were taught by the same instructor. The SI leaders were trained on the SI model, including the use of key facilitation strategies and the development and implementation of SI session plans (Appendix A provides a sample session plan from one of the SI leaders involved with this study). The SI supervisor observed both leaders during their sessions throughout the semester to ensure

they were following the SI model. After session observations, the SI leaders received a

completed feedback form and met with the SI supervisor to discuss their strengths and areas for

improvement (Appendix B provides a sample SI observation record for one of the SI leaders

involved in the study). Throughout the semester, both leaders hosted a combined 6 one-hour

sessions most weeks, with the exceptions of holidays and occasional cancellations (e.g., due to

illness). In all, students in the biology course had the opportunity to attend 69 sessions

throughout the semester.

Measures

This section describes the measures used in the study. I administered 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 collected

SI attendance data from the SI program, exam and final course grades from the course instructor,

and student demographic data from the institutional assessment office.

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 individuals can judge their performance on a task (Bol et al., 2010). In this study, the tasks are the first and final exams taken by students, which were 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 relate to fluctuations in SI attendance and final calibration. To measure beginning calibration, students were asked to predict their grade on the first exam with the following item: "On a scale of 0-100%, predict your grade for this exam" (Serra & DeMarree, 2016). Students selected a response ranging from 0-100 to indicate their predicted exam score. Exams for the course were multiple-choice and were scored using a Scantron device. The course instructor provided students' actual exam scores on a 0-100% scale to measure calibration. Thus, I used global-level (rather than local-level, or item-by-item) judgments for the calibration measure (Hacker et al., 2008). In addition, predictions, rather than postdictions, were used since students tend to be less accurate with predictive judgments (e.g., Bol et al., 2010).

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, this standard was used by calculating the absolute difference between participants' predicted and actual exam scores. Calibration scores ranged from 0 to 92, with lower scores demonstrating greater calibration accuracy and a score of zero indicating perfect calibration. After adjusting for outliers using Grubb's Test (which is explained later in this chapter), the scores ranged from 0 to 69.

I used the absolute differences among predicted and actual scores in the path model. Then, students' bias scores were used to examine the results descriptively. Bias scores were 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 predicted he or she would earn an 80% but received a 50% on the exam, the overconfidence score was +30. Conversely, a student who estimated he or she would produce an 80% but earned a 90% had an underconfidence score of -10. This was important to examine because overconfidence and underconfidence have different implications for learners. Overconfidence could lead to students not investing the appropriate amount of time into studying, while underconfidence may cause students to waste their time by studying easier concepts (Hacker et al., 2008).

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 were 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" (Serra & DeMarree, 2016). Students again selected a response ranging from 0-100 to indicate their predicted scores on the multiple-choice final exam, and the course instructor provided me with students' actual exam scores on a 0-100% scale. Absolute calibration was determined by calculating the differences in their predicted and actual exam scores, and scores ranged from 0 to 91. After adjusting for outliers using Grubb's Test (which is explained later in this chapter), the scores ranged from 0 to 75.

Self-Efficacy Scale

Students' self-efficacy was measured at the beginning and end of the semester when they were asked the exam calibration question. The pre- and post-tests were 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 used an existing scale from the Patterns of Adaptive Learning Scale (PALS) to measure participants' beginning self-efficacy (Midgley et al., 2000). Specifically, students answered the five questions from the PALS Academic Efficacy scale with a minor adjustment of replacing "class" with "biology course" (see Appendix C). Each item asked 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 asked 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 of college students in an undergraduate biology course found that the internal consistency reliability for the Academic Efficacy scale is 0.92 (Perez et al., 2018), 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).

The beginning self-efficacy variable for the path model was calculated by averaging each participant's responses to the five Likert-scale questions. Higher score averages were indicative of higher self-efficacy, and score averages ranged from 1.4 to 5.0.

Final self-efficacy. Final self-efficacy is an endogenous variable that may be predicted by beginning self-efficacy and SI attendance. Prior to the final exam, students again were asked

to respond to the five questions from the PALS Academic Efficacy scale. Responses to the five questions were averaged to produce the final self-efficacy variable, and score averages ranged from 1.0 to 5.0.

SI Attendance

SI attendance is the total number of SI sessions attended by students throughout the fall semester and is represented as a continuous variable ranging from 0 to 40 sessions attended (out of a total 69 sessions offered during the semester). After adjusting for outliers using Grubb's Test (which is explained later in this chapter), the number of sessions attended ranged from 0 to 12. SI leaders collected 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, an Excel report of student SI attendance was collected from the SI program, which was used to match attendance 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 were used in this study. I requested from the course instructor students' final course grades and exam grades (see Appendix D for the request letter sent to the course instructor). Information was also collected 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, and class standing (refer to Appendix E for the letter sent to the institutional assessment office).

Final course grade. The course instructor provided students' final course grades in the general biology course on a 0-100% scale. This was included as an endogenous variable in the hypothesized path model that may have been predicted by the total SAT score, SI attendance, final calibration, and final self-efficacy variables.

Exam grades. Students' grades on the first and final multiple-choice exams also were requested from the course instructor (see Appendix D). These grades were provided on a 0-100% scale and were compared with students' exam grade predictions to calculate beginning and final calibration scores for each student.

Total SAT score. SAT scores were obtained from the institution's assessment office. The scores fit within a range from 400-1600. This is an exogenous variable used in the path model to control for prior achievement.

Other student demographics. Other student demographic variables were obtained from the office of institutional assessment to describe the study participants and the general biology class population. Specifically, characteristics collected include student gender, race/ethnicity, and class standing, e.g., freshman. This information is presented in aggregate form in the next chapter.

Procedure

At the end of the second week of class, one week prior to the first exam, I electronically distributed the student survey using Qualtrics. Students received an email immediately prior to their class time during which I introduced my study and asked students to complete the survey on their electronic devices. The course instructor offered students extra credit for completing the survey or for an alternative assignment of completing problems from the back of the textbook for students who did not wish to complete the survey. Extra credit was removed when calculating

students' final course grades in the path analysis. In addition, students were given the incentive of entering their name into a drawing for one of ten $10 Amazon gift cards, which was awarded to randomly selected students at the end of the semester. Students were 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 included a notification letter informing participants of the study's purpose, requirements, potential benefits and risks, voluntary nature, and assurance of confidentiality. The letter notified them that, should they complete the assessment, their responses would be matched to their demographic characteristics, grades, and SI attendance data. Participants also were notified that the instructor and SI leaders would not have access to survey responses and that their responses would have no effect on their grades (see Appendix F for notification letter). Students had the option to electronically consent to participate in the study prior to answering the survey questions. Students were allotted time during class to complete the survey, and Qualtrics was used to send them reminder emails each day, ending on the day of the exam.

Throughout the semester, SI leaders hosted weekly sessions and asked students to sign-in to the session using an electronic kiosk. In the case of technical difficulties, SI leaders collected student names and university identification numbers (UINs) via paper and entered information retroactively into the electronic system.

One week prior to the final exam, I visited each class section to encourage students to complete the posttest survey emailed to them via Qualtrics immediately prior to their class period. Again, students were 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

provided students with class time to complete the survey, and daily email reminders were sent to students, ending on the date of the final exam.

Once final grades were submitted, I collected students' exam grades and final course grades from the instructor (with extra credit removed). In addition, SI attendance data was collected from the electronic system, and additional student performance and demographic data were requested from the institutional assessment office. Students' UINs were used to merge all records, and IBM SPSS Statistics 24 and Mplus (v 7.3) were used for all data analysis.

I ensured confidentiality by asking students to use their UINs when completing both the pre- and post-assessments. In addition, participant information was 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 conducted once data was collected. First, I describe the descriptive statistics. Then, I explain the process of checking for assumptions and how I conducted my path analysis.

Descriptive Statistics

I began my data analysis, described in the next chapter, by examining the descriptive statistics of my collected data. The first set of data involved calculating frequencies and percentages of demographic factors, including gender, race/ethnicity, year, and class standing to assess the representativeness of the sample to the larger population of the general biology course. I also obtained the means, standard deviation, skew, and kurtosis for the path model's seven variables. As expected, SI attendance was not normally distributed, so the frequencies and

percentages for SI attendance data were detailed. Finally, a correlation matrix of the study

variables is provided in Chapter Four.

Checking for Assumptions

Prior to applying my path analysis, my data was examined to make sure it fit with

common assumptions used for multivariate procedures. According to Keith (2015), there are

five major assumptions that underline the use of multiple regression and path modeling,

including: (a) linearity, (b) independence of observations, (c) homoscedasticity, (d)

multicollinearity, and (e) normality. In addition, Keith recommends using distance, leverage,

and influence to diagnose data problems referred to as "outliers" or "extreme cases" (p. 195).

The tests recommended by Keith (2015) were used to verify that the first four

assumptions were met. First, the curve estimation feature was utilized in SPSS (v 24) to verify

that the assumption of linearity of the data was not violated for linear regressions of final course

grade on SI attendance, final calibration, and final self-efficacy. Next, the assumption of

independence of observations was met because participants in the study were enrolled in a class

with the same professor and were administered the same pre- and post-test measures with no

observers required. Third, the assumption of homoscedasticity was met by inspecting the shape

of the data in the scatterplots of the residuals from the regressions referenced above. Fourth, I

requested collinearity diagnostics from SPSS to test for multicollinearity and discovered that

Tolerance scores were all close to 1 and variance inflation factor (VIF) scores were all well

below 6, meeting the desirable score ranges outlined by Keith (2015).

Unlike the first four assumptions, non-normality and outliers were assumptions not met

by the dataset. According to Byrne (2012), a critically important assumption of a path analysis is

that the data are multivariate normal, and data that are multivariate kurtotic are especially

problematic to path model analyses. As suspected, this was an issue with my dataset, especially

for the SI attendance variable. Robust maximum likelihood estimation was used in Mplus (v

7.3) to account for the non-normal distribution of the SI attendance data (Byrne, 2012). In

addition, Grubb's Test was used to identify and adjust for outliers (Grubbs, 1969). Instead of

completely removing the outliers that were discovered for several of the variables (total SAT,

beginning and final calibration, SI attendance, and final course grade), the critical values of z

found for each variable in the Grubb's Test were used to change the outliers to the next highest

or lowest variables within the dataset. I used these adjusted variables in the path model analysis.

Path Analysis

I applied a path analysis with robust maximum likelihood (MLR) estimation to answer

the research questions using Mplus (v 7.3; Byrne, 2012). MLR was used to account for the

nonnormally distributed data, as well as the incomplete data for the total SAT score variable.

Again, the hypothesized path model is in Figure 2. The cutoff value for statistically significant

results was $p < .05$.

After running the analyses for the hypothesized model, fit statistics recommended by Hu

and Bentler (1999) were used to assess model fit, beginning with chi-square (X^2). A path model

is considered a good fit if the chi-square statistic is small and non-significant. Due to the

sensitivity of X^2 to sample size, other indicators of model fit were used, including the

Comparative Fit Index (CFI), 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 CFI greater than .95,

RMSEA less than .06, and SRMR of less than .05 (Byrne, 2012; Hu & Bentler, 1999).

While the hypothesized path model was based on theoretical and empirical literature, the original model was rejected due to its poor fit with the sample data (Byrne, 2012). Specifically, the X^2 statistic was very high and significant, and the CFI, RMSEA, and SRMR statistics did not meet the cutoff values recommended by Byrne (2012) and Hu and Bentler (1999).

As a result, I engaged in a process known as "model generating" (Byrne, 2012, p. 8) by which the modification indices (modindices) were used from the original path model to identify and determine statistically significant improvements to develop an adjusted path model (Loehlin, 1998). The modindices indicated that beginning calibration was a significant predictor of final self-efficacy and final course grade, which also makes sense theoretically, so these paths were added in the adjusted path model.

The resulting path model is in Figure 3. When displaying the adjusted path model, several non-significant paths were removed, including the correlational arrows between total SAT score and beginning self-efficacy and between final calibration and final self-efficacy. In addition, the depicted adjusted path model reflects paths removed from beginning calibration to SI attendance and from SI attendance to both final calibration and final self-efficacy.
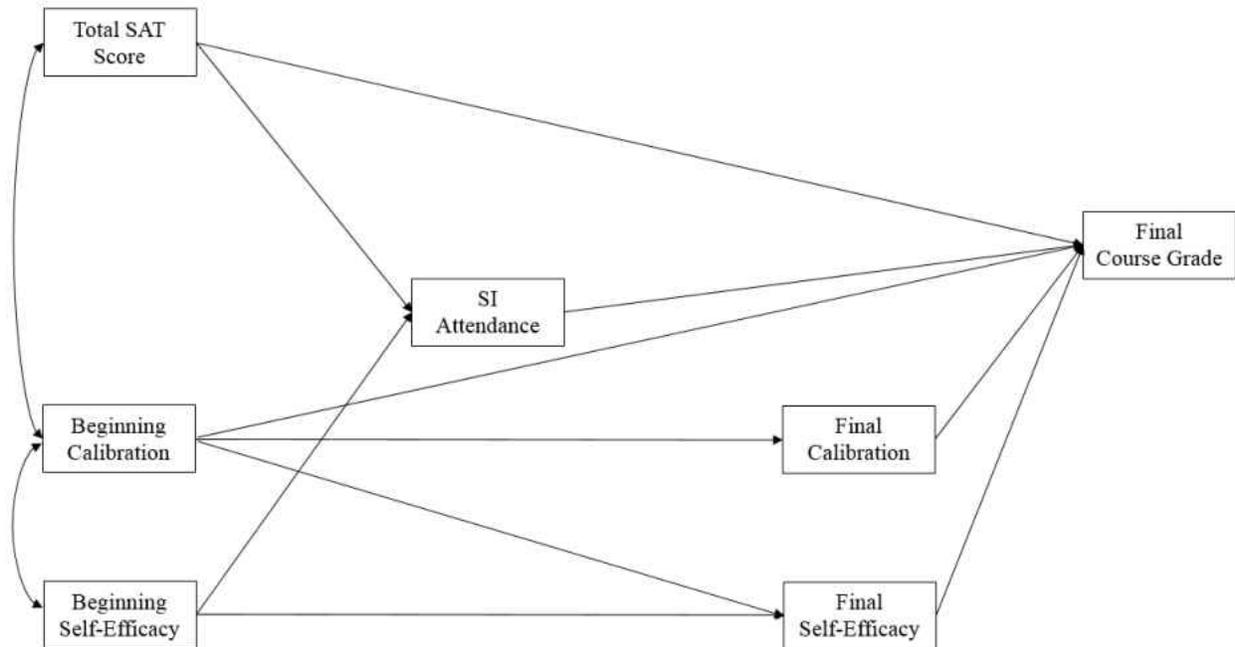
Figure 3. Adjusted path model after model generating process. Only significant paths are presented for simplicity.

The adjusted path model resulted in a chi-square that was still significant and a little high, but it was much improved from the original path model. In addition, the fit statistics for the final path model included a CFI of .96, RMSEA of .12, and SRMR of .05. While the RMSEA statistic was higher than the ideal cutoff value of .06, the CFI and SRMR fit within the recommended ranges.

Summary

In this chapter, I described the methodology used to address the research questions and hypotheses that were shared at the beginning of the chapter. I presented my hypothesized path model and study participants, including their characteristics, university context, and the SI program that has been certified by the International Center for SI. In addition, information was provided for the path model variables, including the exogenous variables of total SAT score,

beginning calibration, and beginning self-efficacy and endogenous variables of SI attendance, final calibration, final self-efficacy, and final course grade. The procedures used to collect this data were also described in detail. Finally, I provided an overview of the data analysis, including the descriptive statistics, steps for checking for and addressing violations of assumptions, and process of using Mplus to apply the path analysis and arrive at the adjusted model presented above in Figure 3. Next, the fourth chapter describes the findings generated from these methodological procedures.

CHAPTER FOUR

FINDINGS

The third chapter described the study methodology, including the research questions and hypotheses, study design, participants and context, measures, procedure, and data analysis. In this chapter, I present the results of the study. First, I provide descriptive statistics for the population and participant characteristics and the statistics and correlations for the path model variables. Then, I share the path model findings that address the three research questions, as well as other observations regarding the relationships among the variables.

Descriptive Statistics

In this section, descriptive statistics for the variables are shared. Specifically, I provide the demographic characteristics for students from the entire class population and the participants from the study sample. Then, descriptive statistics are presented for the variables in the path model, including mean, standard deviation, skew, and kurtosis, as well as the frequencies and percentages for the non-normal SI attendance variable. Finally, I provide a correlation matrix to show the relationships among the variables in the path model.

Population and Participant Characteristics

Table 4 provides an overview of the characteristics of the study participants and the students enrolled in the general biology course at the end of the term. Student characteristics (i.e., gender, race/ethnicity, and class standing) were provided by the institution's assessment office. In addition, SI participation rates were collected from the SI Program supervisor. Students' University Identification Numbers (UINs) were used to match all student data.

Table 4

Characteristics of General Biology Students from the Class Population and Study Participants at the End of Term

Characteristic	Class Population (*n* = 482)		Study Participants (*n* = 320)	
	n	*p*	*n*	*p*
Gender				
Female	327	67.8	230	71.9
Male	155	32.2	90	28.1
Race/Ethnicity				
African-American	172	35.7	115	35.9
Caucasian	171	35.5	112	35.0
Two or more races	49	10.2	31	9.7
Hispanic/Latino	41	8.5	28	8.8
Asian	30	6.2	21	6.6
Native Hawaiian/Pacific Islander	1	0.2	0	0.0
Non-resident aliens	3	0.6	2	0.6
Unknown	15	3.1	11	3.4
Class Standing				
Freshman	267	55.4	195	60.9
Sophomore	134	27.8	76	23.8
Junior	55	11.4	36	11.3
Senior	16	3.3	10	3.1
Graduate/Unclassified	10	2.1	3	0.9
Attended SI 1+ times	124	25.7	93	29.1

Note. Demographic data is reported for students still enrolled in the general biology course at the end of the semester.

Chi-square tests of independence were used to compare the study participants and class population in terms of gender, race/ethnicity, and class standing. There were two areas in which the sample differed significantly from the class population: gender and class standing. There were more females who participated in the study when compared with the class population, X^2 (1, $N = 482$) = 7.10, $p = .01$. Since the class standing variable includes five categories, four dummy variables were created with Freshman serving as the reference group in each dummy variable, as recommended by Keith (2015). The four dummy variables were Sophomore, Junior,

Senior, and Non-Degree. Chi-square tests of independence for the four class standing variables revealed a significant difference in the Freshman vs. Sophomore variable between the sample, X^2 $(1, N = 482) = 7.78, p = .01$. More freshmen than sophomores participated in the study when compared with the class population. In addition, an independent samples t-test revealed a significant difference in SI attendance rates between the sample $(M = 1.21, SD = 4.10)$ and class population $(M = .46, SD = 1.40)$; $t(480)=2.52, p=.03$.

Path Model Descriptive Statistics

Descriptive statistics for all the variables used in the path model are presented in Table 5, including the number of cases, mean, standard deviation, skewness, and kurtosis for each variable. Inspections of histograms and stem-and-leaf plots revealed normal distributions for the total SAT score and final course grade variables, while the other variables were not normally distributed. Beginning and final calibration and SI attendance variables were skewed to the left, indicating more instances of lower scores. Conversely, beginning and final self-efficacy variables were skewed to the right, indicating more instances of higher scores. In addition, the kurtosis scores for SI attendance and final calibration were above the conservative range of ± 2.0 recommended by Byrne (2012). To account for this, a maximum likelihood estimation was used in the path model analysis (Byrne, 2012).

Table 5

Descriptive Statistics for Path Model Variables

Variable	*n*	*M(SD)*	Skewness	Kurtosis
Total SAT Score	257	1053.81 (171.59)	-0.84	1.99
Beginning Calibration	320	19.23 (15.34)	0.92	0.41
Beginning Self-Efficacy	320	4.11 (.75)	-1.15	1.45
SI Attendance	320	0.92 (2.25)	3.56	13.47
Final Calibration	320	15.97 (15.37)	1.74	3.39
Final Self-Efficacy	320	3.89 (.9)	-0.53	-0.41
Final Course Grade	320	71.44 (14.28)	-0.42	-0.14

Note. Descriptive statistics are provided for the variables after adjusting for extreme outliers in the total SAT, beginning and final calibration, SI attendance, and final course grade variables.

The SI attendance variable is unique in its skewness and kurtosis due to the large number of students who attend zero sessions during the semester. To provide more context for this variable, Table 6 presents the frequencies and percentages for the number of students who attended zero to 12 sessions throughout the term. This table reports the seven outliers as having attended 12 sessions each, which was the next highest variable when using the critical value of z in the Grubb's Test (Grubbs, 1969). The outliers actually attended 17, 19, 19, 23, 29, 31, and 40 sessions.

Table 6

SI Attendance Frequencies and Percentages

No. of SI Sessions Attended	f	p
0 sessions	227	70.9
1 session	41	12.8
2 sessions	17	5.3
3 sessions	12	3.8
4 sessions	3	.9
5 sessions	4	1.3
6 sessions	4	1.3
7 sessions	2	.6
8 sessions	2	.6
12 sessions	8	2.5

Note. The table represents the adjusted SI attendance variable used in the path model. Seven outlier variables of 17, 19, 19, 23, 29, 31, and 40 sessions were reconfigured to 12 sessions using the Grubb's Test critical value of *z*.

Path Model Variable Correlations

Table 7 outlines the bivariate correlations calculated in SPSS (v 24). This analysis demonstrates that there were both strong and weak relationships among the variables.

Table 7

Path Model Variable Correlations

	1	2	3	4	5	6	7
1. Total SAT Score	–						
2. Beginning Calibration	-.40**	–					
3. Beginning Self-Efficacy	.10	-.12*	–				
4. SI Attendance	-.15*	-.04	-.14*	–			
5. Final Calibration	-.29**	.31**	-.09	-.06	–		
6. Final Self-Efficacy	.24**	-.39**	.42**	-.02	-.12*	–	
7. Final Course Grade	.40**	-.65**	.22**	.13*	-.57**	.55**	–

Note. *p < .05, **p < .01

Looking at the correlations among the variables, final course grade had a statistically significant relationship with all the variables in the path model. This variable was most associated with beginning calibration, $r(320) = -.65$, $p < .001$ and final calibration, $r(320) = -.57$, $p < .001$. In other words, students who were more accurate in predicting their test scores at the beginning and end of the term were more likely to have higher final course grades. Final course grade also had a significant relationship with final self-efficacy, $r(320) = .55$, $p < .001$; whereas, the relationship between final course grade and beginning self-efficacy was significant but not as strong, $r(320) = .22$, $p < .001$. In addition, as expected, students with higher total SAT scores were more likely to have higher final course grades, $r(257) = .40$, $p < .001$. There also was a relationship between SI attendance and final course grade, but SI attendance had the weakest correlation with final course grade among all of the path model variables, $r(320) = .13$, $p = .02$.

Total SAT score is another variable that had a significant relationship with most of the variables, though it should be noted that 19.7% of the participants were missing SAT data due to the site not requiring this information for admission. The only exception was that there was no statistically significant relationship between total SAT score and beginning self-efficacy, which is interesting. One would expect that students with higher standardized test scores would have a higher sense of self-efficacy at the beginning of a challenging college-level course; however, this was not true for the sample population.

Another observation from the variable correlations was that students' self-efficacy at the beginning of the semester was positively correlated with their self-efficacy at the end of the term, $r(320) = .42$, $p < .001$. In addition, students' beginning calibration accuracy was positively correlated with their final calibration, $r(320) = .31$, $p < .001$. One final relationship to note is that students' beginning self-efficacy was positively correlated with their beginning calibration

accuracy, $r(320) = -.12$, $p = .03$. The same was true for the relationship between final self-efficacy and calibration, $r(320) = -.12$, $p = .03$.

Primary Analysis

Results of the path model analysis that address the three primary research questions are presented next. Figure 4 shows the final model with only significant paths displayed and the standardized direct effects of the exogenous (or independent) variables on the endogenous (or dependent) variables. The Chi-square statistic in the adjusted path model was relatively large: X^2 $(4, N = 320) = 21.93$, $p < .001$; however, it was much improved from the original path model. In addition, the root mean square error of approximation (RMSEA) = .12 was a little high. The other indices indicated a good model fit: comparative fit index (CFI) = .96 and standardized root mean square residual (SRMR) = .05 (Byrne, 2012; Hu & Bentler, 1999).

The first research question explored how students' self-regulated learning behaviors at the beginning of the semester influenced their decisions to attend SI. The second and third research questions examined how SI attendance may have directly influenced students' final calibration, self-efficacy, and course grades, as well as how SI attendance may have indirectly influenced final course grades through calibration and self-efficacy. Results for research questions two and three are presented together.
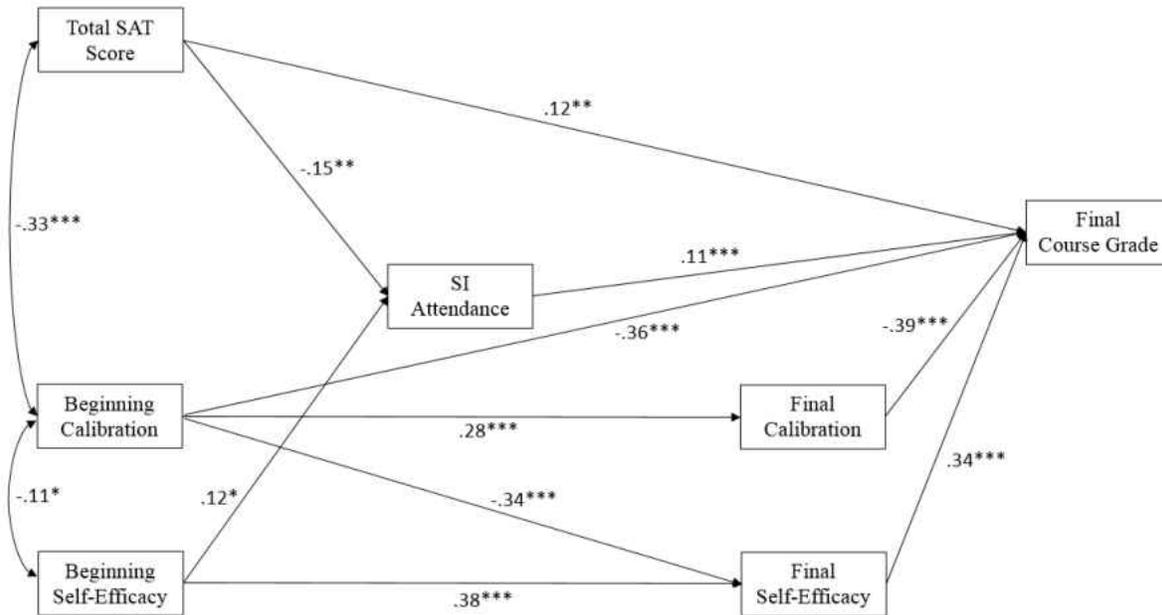
Figure 4. Adjusted path model results. Only significant paths are displayed for simplicity. Standardized coefficients are presented; *p* < .05, **p* < .01, ***p* < .001

RQ1: Beginning Self-Efficacy and Calibration as a Predictor of SI Attendance

After controlling for total SAT score and beginning calibration, students' beginning self-efficacy predicted their SI attendance during the semester (β = -.12, p = .03). In other words, students with lower self-efficacy were more likely to participate in SI.

Conversely, results indicated that beginning calibration accuracy did not predict SI attendance (β = -.09, p = .17). Thus, students' ability to predict their first exam scores did not influence their decision to attend SI. The path model explained only 4% of the total variance in SI attendance (R^2 = .038) at the beginning of the semester.

Since the absolute differences among predicted and actual scores were used for beginning calibration in the path model, students' bias scores were used to examine these results descriptively. Positive bias score numbers were reflective of students being overconfident, while

negative numbers represented underconfidence (Hawthorne, et al., 2017). To examine the

differences between over- and under-confident beginning calibration responses, students were

divided into two groups. Three students were excluded from this grouping based on their perfect

prediction scores of zero. An independent samples t-tests revealed no statistically significant

differences in SI attendance patterns between those who were initially over-confident ($M = .80$,

$SD = 2.09$) and those who were initially under-confident ($M = 1.28$, $SD = 2.63$); $t(306)=-1.42$,

$p=.16$. This finding further indicates the lack of influence calibration accuracy has on SI

attendance patterns among the participants.

RQ2 and RQ3: SI Attendance as a Direct and Indirect Predictor of Final Calibration, Self-Efficacy, and Course Grades

In addition to exploring if calibration and self-efficacy at the beginning of the semester

predicted SI attendance, I also examined if SI attendance predicted students' calibration and self-

efficacy at the end of the term. In addition, I examined whether SI attendance had a direct

influence on final course grade, as well as an indirect effect on final course grade through final

calibration and self-efficacy.

Results indicated that SI attendance did not predict final self-efficacy ($\beta = .35$, $p = .36$),

after controlling for total SAT score, beginning self-efficacy, and beginning calibration. Thus,

attending SI did not directly improve or worsen students' self-efficacy at the end of the semester

(RQ2). In addition, there was no indirect effect of SI attendance on final course grade through

final self-efficacy (RQ3).

After controlling for total SAT score, beginning calibration, and beginning self-efficacy,

the results of the path model analysis also indicated that SI attendance did not predict final

calibration ($\beta = -.04$, $p = .57$). This means that students' participation in SI had no direct

influence on students' calibration accuracy at the end of the semester (RQ2). Furthermore, SI did not indirectly predict students' final course grades through final calibration (RQ3).

Once again, students' bias scores were used for their final calibration to descriptively analyze whether there were any differences between those who over- and under-predicted their performance on the final exam. An independent samples t-test revealed statistically significant differences in SI attendance patterns between those who were over-confident at the end of the semester ($M = .76$, SD $= 1.93$) and those who were under-confident at the term's conclusion ($M = 1.58$, SD $= 3.17$); $t(310)=-2.66$, $p=.01$. This analysis indicates that students who attended SI more frequently were more likely to underpredict how well they would perform on the final exam, while those who attended SI less often had a more inflated view their final exam scores.

Finally, after controlling for total SAT score, beginning calibration, and beginning self-efficacy, SI attendance predicted students' final course grades ($\beta = .11$, $p < .001$). In other words, SI attendance had a small, direct, and significant effect on students' final course grades, meaning that students who attended SI more frequently performed better in the course (RQ2). Furthermore, it is valuable to use the unstandardized coefficient for final course grade on SI attendance ($b = .38$, $p < .001$) to interpret this data. This measure indicates that, for each SI session attended, participants' grades increased by .38 points on a 1 to 100 scale. In other words, students could increase their grade by one percentage point after attending three SI sessions or by approximately half a letter grade by attending 13 SI sessions (or once/week during the semester). However, as the results have indicated, while SI had a direct effect on final course grade, there was no indirect effect of SI attendance on final course grade through final calibration or self-efficacy (RQ3). This means that there must be other factors from attending SI that influenced

students' final course grades. Overall, the path model explained 67% of the total variance in final course grade ($R^2 = .673$) at the end of the semester.

Using the data available, the potential influence of SI attendance on students' performance in the course was further explored by dividing students into two groups based on if they improved from the first to the final exam or if their grades worsened from test one to test five. An independent samples *t*-test revealed statistically significant differences in SI attendance patterns between those who improved their tests scores from the beginning to the end of the semester (*M* = 1.22, SD=2.62) and those who had worse grades on their final than on their initial exam (*M* =.50, SD=1.58); $t(299)=-2.94$, *p*=.004. This further supported the finding that SI attendance was positively correlated with students performing well at the end of the semester, specifically on their final exams when compared with their first test scores.

Other Findings

The results of the final adjusted path model revealed several findings that fall outside of the scope of the research questions for the study, but they are nonetheless relevant to the overall purpose of the study. In this section, I discuss these findings, first focusing on the exogenous (or independent) variables of total SAT score, beginning calibration, and beginning self-efficacy. Then, I shift the discussion to the two endogenous (or dependent) variables with notable relationships: final self-efficacy and final calibration.

Exogenous Variables

In the original path model, all three of the exogenous variables of total SAT score, beginning calibration, and beginning self-efficacy were predicted to be correlated with one another. As expected, beginning self-efficacy and beginning calibration were correlated (*r* = -.11, *p* = .048). In other words, students with higher self-efficacy were more likely to calibrate

accurately their exam scores at the beginning of the semester. In addition, beginning calibration

was correlated with total SAT score (r = -.33, p < .001), indicating that students with higher SAT

scores were more likely to accurately predict their first exam scores. A surprising finding among

the exogenous variables was that beginning self-efficacy was not correlated with total SAT score

(r = .09, p = .12). One would think that students with higher SAT scores would enter

challenging college courses with higher self-efficacy; however, this was not the case among this

sample of general biology students.

In addition, as predicted, total SAT score was a significant predictor of SI attendance and

final course grade. After controlling for beginning calibration and self-efficacy, students' total

SAT scores predicted their SI attendance during the semester (β = -.15, p = .002). This makes

sense in that students who knew that they were not high performers based on prior achievement

were more likely to attend SI, acknowledging that they would likely benefit from the

intervention. However, this finding is puzzling when one considers that total SAT score was not

correlated with beginning self-efficacy. In other words, low SAT scores appeared to motivate

students to attend SI for additional help; however, SAT scores did not affect their self-reported

beginning self-efficacy. This could be because SAT is more of a global measure, while self-

efficacy was for the course specifically. In addition, after controlling for all the study's

variables, total SAT score predicted final course grade (β = .12, p = .004). Thus, students with

higher prior achievement were more likely to earn higher final course grades, which was an

expected finding. Once again, the path model explained only 4% of the total variance in SI

attendance (R^2 = .038) at the beginning of the semester and 67% of the total variance in final

course grade (R^2 = .673) at the end of the semester.

An unexpected finding among the exogenous variables in the study was the significant influence of beginning calibration on two endogenous variables: final course grade and final self-efficacy. After observing the modindices for the original path model, these were identified as significant paths to add in the adjusted path model. After controlling for total SAT score and beginning self-efficacy, students' beginning calibration predicted their final self-efficacy (β = -.34, p < .001). In other words, students with better prediction accuracy at the beginning of the term had higher self-efficacy at the end of the semester. In addition, there was a stronger relationship between beginning calibration and final self-efficacy than there was between beginning and final calibration (β = .28, p < .001). An additional unexpected finding after observing the modindices of the original path model was that beginning calibration significantly predicted final course grade after controlling for other variables in the model (β = -.36, p < .001). Therefore, students with better calibration accuracy at the beginning of the semester were more likely to end the semester with higher final course grades. As I will present in the next section, beginning calibration was almost as strong of a predictor of final course grade as final calibration, which makes sense when considering the statistically significant relationship between beginning and final calibration. The path model explained 29% of the variance in final self-efficacy (R^2 = .286), 8% of the variance in final calibration (R^2 = .083), and 67% of the total variance in final course grade (R^2 = .673) at the end of the semester.

A final and expected finding among the exogenous variables was the relationship between beginning and final self-efficacy. After controlling for total SAT and beginning calibration, beginning self-efficacy predicted final self-efficacy (β = .38, p < .001). Therefore, as with calibration, students' self-efficacy at the beginning of the semester predicted their final self-efficacy.

Endogenous Variables

In this final section, I provide observations of the relationships in the path model among the endogenous (or dependent) variables that fall outside of the research questions. First, unlike students' beginning self-efficacy and calibration, their final self-efficacy and calibration were not correlated ($r = .02$, $p = .68$). This finding was unexpected since there was a significant relationship between these variables at the start of the semester and beginning calibration and self-efficacy predicted final calibration and self-efficacy, respectively.

In addition, the path model revealed expected findings for the relationship between final calibration and final course grade. After controlling for total SAT score, beginning calibration and self-efficacy, and SI attendance, final calibration predicted final course grade ($\beta = -.39$, $p < .001$). Again, the overall path model explained 67% of the total variance in final course grade ($R^2 = .673$). This means that students with strong calibration accuracy at the end of the semester earned higher final grades in the course. While this finding makes sense, the results of the path model indicated that this relationship was not because of the influence of SI attendance on final calibration, as expected in the third research question. To explore the differences between students who over- and under-predicted their final exam performances at the end of the semester, an additional analysis was conducted. An independent samples t-test revealed statistically significant differences in final course grades between those who over-predicted their final exam scores ($M = 68.74$, SD $= 13.98$) and those who were under-confident at the end of the term ($M = 81.08$, SD $= 10.30$); $t(310)=-6.74$, $p<.001$. So, while the path model results indicated that there was a strong relationship between final calibration accuracy and final course grades, this especially was accurate for students who tended to under-predict their final exam scores.

Finally, after controlling for total SAT score, beginning calibration and self-efficacy, and SI attendance, final self-efficacy also predicted final course grade ($\beta = .34$, $p < .001$). So, students with higher self-efficacy at the end of the semester tended to earn higher final course grades. Again, this finding is intuitive, though it was not due to the influence of SI attendance on students' self-efficacy as predicted in the original path model. To dig deeper into this relationship between final self-efficacy and final course grade, a regression of final course grade on the change in students' beginning self-efficacy to final self-efficacy was conducted to reveal a significant 14% of the variance in final course grade; $F(1, 318) = 49.81$, $MSE = 176.75$, $p < .001$. In other words, students who saw an increase in self-efficacy from the beginning to the end of the semester also experienced significant increases in their final course grade.

Summary

In this chapter, I provided the findings of the path analysis to examine the influence of self-efficacy and calibration on SI attendance, as well as the relationship between SI attendance and final self-efficacy, calibration, and course grades. This was done by first comparing the characteristics of the study participants to the population of the biology courses. While the sample was comparable to the class population in terms of race/ethnicity, the participants were more heavily represented by females, freshmen, and more frequent SI attendees. Then, I shared the descriptive statistics for the path model variables, including the number of cases, mean, standard deviation, skewness, and kurtosis for each variable. A maximum likelihood estimation was used in the path model analysis to account for the abnormal skewness and kurtosis for many of the variables, especially SI attendance. Next, correlations for the path model variables were presented to demonstrate strong relationships among the variables. Of note was the statistically significant correlation that final course grade had with all the variables in the path model.

After providing the preliminary descriptive statistics, the path analysis results were presented. Specifically, the findings addressed the first research question by indicating that self-efficacy at the beginning of the semester was a significant predictor of SI attendance; however, beginning calibration was not a significant predictor of SI participation. The path model also indicated that SI attendance was a significant predictor of final course grade, though it was not a predictor of final calibration or self-efficacy (RQ2). The results of the second research question also addressed RQ3, indicating that there was not an indirect effect of SI attendance on final course grade through final calibration and self-efficacy. After addressing the research questions, other findings were shared from the adjusted path model, including the unexpected and significant influence of beginning calibration on final self-efficacy and final course grade and the anticipated finding that final self-efficacy and calibration predicted final course grade. Chapter 5 will provide a summary of the study along with a discussion of the findings in the context of the previous literature. This chapter also will discuss limitations of the study and implications for future research and practice.

CHAPTER FIVE

DISCUSSION

This chapter summarizes the major findings from this study of the connections between a

Supplemental Instruction program and the self-regulated learning constructs of self-efficacy and

calibration. The primary focus of this study was to investigate if students' pre-existing self-

efficacy beliefs and calibration accuracy predicted their decisions to attend SI sessions

throughout the semester. In addition, the study explored if SI attendance had a direct effect on

changes in students' self-efficacy and calibration and subsequent indirect effects on students'

final course grades. While previous research has looked the relationship between self-efficacy

and SI, this study attempted to account for prior methodological limitations. In addition, this was

the first known study to examine calibration accuracy and its association to any academic

support program. The exogenous (or independent) variables of total SAT score, beginning

calibration, and beginning self-efficacy were studied for their effects on the endogenous (or

dependent) variables of SI attendance, final calibration, final self-efficacy, and final course

grade. This study employed a path model analysis with robust maximum likelihood estimation

using Mplus (v 7.3; Byrne, 2012) and fit criteria recommended by Hu and Bentler (1999). The

path analysis answered the following research questions:

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?

3. What is the indirect effect of SI attendance on final course grades through calibration and self-efficacy?

Summary of Results

In this section, I present a summary of the findings presented in Chapter Four. I first address the findings of the three research questions and then provide other observations from the path model variables' relationships.

The first research question examined the extent to which students' self-efficacy beliefs and calibration accuracy at the beginning of a general biology course predicted their SI attendance during the semester. The results of the path model analysis indicated that participants with lower self-efficacy were more likely to participate in SI; however, students' calibration accuracy at the beginning of the semester did not predict their decision to attend SI. Furthermore, there was no difference in SI attendance between participants who were over- and under-confident in their first exam score predictions. The hypotheses were neither supported nor refuted since there were conflicting views in the literature on the influence of self-efficacy on SI participation or help seeking. In addition, there was no prior research on how students' calibration accuracy may influence their participation in academic support programs.

The second research question explored the direct influence of SI attendance on final self-efficacy, calibration accuracy, and course grades at the end of a general biology course. Findings of the study indicated that SI attendance did not directly predict final self-efficacy or calibration accuracy. However, SI attendance did have a modest, direct effect on participants' final course grades, meaning that students who attended SI more frequently performed better in the course. Thus, the hypothesis that SI would positively influence final calibration accuracy was not supported, but my prediction that SI would positively affect final course grade was supported. In

addition, conflicting research made the predicted relationship between SI attendance and self-efficacy unclear, and the path model revealed that there was no significant relationship among these variables.

The third research question addressed the potential indirect effects of SI attendance on final course grades through final calibration and self-efficacy. The lack of significant relationship between SI attendance and final self-efficacy and calibration in the path model indicated that, while SI had a direct effect on final course grade, there was no indirect effect of SI attendance on final course grade through final self-efficacy or calibration. Overall, the path model explained 67% of the total variance in final course grade at the end of the semester.

The path model revealed other significant relationships among the variables. Specifically, the path model analysis revealed statistically significant relationships among the exogenous variables of beginning self-efficacy and calibration and between beginning calibration and total SAT score; however, there was no relationship between beginning self-efficacy and total SAT score. These findings indicate that (a) students with higher self-efficacy were more likely to calibrate accurately their exam scores at the beginning of the semester, (b) participants with higher SAT scores were more likely to predict their first exam scores accurately, and (c) total SAT scores did not influence students' self-efficacy at the beginning of the general biology course.

In addition, participants' total SAT scores predicted their SI attendance during the semester and their final course grade. Specifically, students with lower SAT scores were more likely to attend SI, and students with higher SAT scores were more likely to earn higher final course grades.

The path model results also revealed statistically significant relationships among beginning calibration and final course grade and final self-efficacy. Participants with better calibration accuracy at the beginning of the semester were more likely to end the semester with higher final course grades. In addition, beginning calibration was almost as strong of a predictor of final course grade as final calibration. Students with better prediction accuracy at the beginning of the term also had higher self-efficacy at the end of the semester. In fact, there was a stronger relationship between beginning calibration and final self-efficacy than there was between beginning and final calibration. Finally, among the exogenous variables, beginning calibration predicted final calibration, and beginning self-efficacy predicted final self-efficacy.

The path model also revealed significant relationships among the endogenous variables. First, unlike participants' beginning self-efficacy and calibration, final self-efficacy and calibration were not correlated. In addition, final calibration predicted final course grade revealing that students with strong calibration accuracy at the end of the semester earned higher final grades in the course. An independent samples *t*-test revealed statistically significant differences in final course grades between those who were over- and under-confident at the end of the term. This indicated that the strong relationship between final calibration and course grades was especially true for participants who were underconfident in their final exam score predictions.

Final self-efficacy also predicted final course grade, demonstrating that students with higher self-efficacy at the end of the semester earned higher final course grades. An additional linear regression of final course grade on the change in participants' beginning and final self-efficacy revealed that students who had an increase in self-efficacy from the beginning to the end of the semester also experienced significant increases in their final course grade.

Discussion of the Research Findings

Now a summary of the research findings has been presented, study results will be discussed in the context of the literature on Supplemental Instruction, calibration, and self-efficacy. First, I will discuss the findings related to the three research questions. Then, the other results of the study will be addressed.

Beginning Self-Efficacy and Calibration and SI Attendance

The first research question addressed whether students' initial self-efficacy beliefs and calibration accuracy influenced their SI attendance. The path model revealed that students with lower beginning self-efficacy were more likely to attend SI. This finding supports the small number of studies that previously have examined this phenomenon (Hizer, 2010; McGee, 2005). In addition, the study revealed no statistically significant correlation between beginning calibration and SI attendance. No previous literature has looked at the influence of calibration on students' participation in SI or related academic support programs.

Beginning self-efficacy influences SI attendance. Study participants with lower self-efficacy were more likely to attend SI. Most of the research conducted on SI and self-efficacy has focused on the influence of SI attendance on self-efficacy at the end of the semester, with the exceptions of McGee (2005) and Hizer (2010) who examined initial self-efficacy and SI attendance. While the influence of SI on final self-efficacy was addressed in the second and third research questions, this study also examined whether students' beginning self-efficacy in a biology course predicted their SI attendance patterns during the semester.

Looking at the help-seeking literature, Newman (2008) suggested from a theoretical perspective that students with high self-efficacy and the ability to predict their need for help would participate in an academic support intervention, like SI, if they determined it was needed.

Conversely, these students would not attend SI if they did not determine that it was needed. The author came to this conclusion by relating students' help-seeking behaviors and self-efficacy to their adaptive and non-adaptive actions as well as their performance and mastery goal orientations.

However, two previous studies that explored the relationship between beginning self-efficacy and SI attendance indicated that there was a negative relationship between initial self-efficacy and SI attendance. 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 SI-supported humanities and science courses, including biology, at a large state university. Dividing the participants into three groups of (a) non-participants, (b) a high-engagement group, and (c) a low-engagement group, the researcher found a statistically significant negative correlation between student participation in SI and their self-efficacy scales. This finding revealed that students with lower self-efficacy at the beginning of the term were more likely to participate in SI. Similarly, Hizer (2010) found that students who attended five or more SI sessions had initially higher levels of anxiety than students who attended zero to four sessions. This research was conducted on 248 students in SI-supported science courses, including biology, at a small, public, four-year university.

Based on the findings of McGee (2005) and Hizer's (2010) studies, one would hypothesize that students with low self-efficacy would be more likely to participate in SI than those with high self-efficacy. However, Newman's (2008) emphasis on the role of adaptive help-seeking behaviors and mastery goal orientations theorized that students' participation in SI may be more influenced by their ability to determine if seeking help was necessary. Thus, prior to conducting the study, it was unclear if self-efficacy would be correlated positively or

negatively with SI attendance. After examining the results of the path model, the previous findings by McGee (2005) and Hizer (2010) were supported because participants in this study with lower beginning self-efficacy attended SI more frequently than their peers did.

       Beginning calibration does not influence SI attendance. Conversely, there is no existing research on calibration and SI or on calibration and help-seeking behaviors. Therefore, it was unclear if calibration accuracy would predict students' SI attendance. The results of the path analysis indicated that students' beginning calibration accuracy did not influence their decision to participate in SI, and this was true for students who both under- and over-predicted their first exam scores.

SI Attendance and Final Calibration, Self-Efficacy, and Course Grades

       The second research question explored the influence of SI attendance on students' final calibration accuracy, self-efficacy, and course grades. The third research question was related in its exploration of potential indirect effects of SI attendance on final course grade through final calibration and self-efficacy. The path model results revealed that SI attendance did not influence final calibration or self-efficacy; however, there was a significant relationship between SI attendance and increased final course grades, as predicted. Theoretical connections to the calibration literature had indicated that SI attendance may increase students' final calibration accuracy; however, this connection was not supported by the findings. It was less clear if SI attendance would influence final self-efficacy due to conflicting findings in the literature. The positive correlation between SI and final course grades found in this study is widely supported by the SI literature. In this section, I contextualize my findings within the literature on SI, calibration, and self-efficacy.

SI attendance does not influence final calibration. The findings of the path model

analysis did not support my hypothesis: SI attendance did not predict final calibration accuracy

for the general biology students from the sample. While no one has studied SI participation and

its influence on calibration, theoretical connections from calibration research address the

potential relationship between the two variables. Hacker and Bol (2019) argued that research

interventions that targeted all three phases of Zimmerman's (2000, 2002) SRL model were more

likely to improve participants' calibration accuracy and academic performance. Specifically,

Bol, Hacker, Walck, and Nunnery (2012) targeted the SRL model's forethought, performance,

and self-reflection phases in their individual and group guidelines 2 x 2 quasi-experimental study

on high school biology students. The researchers found that participants who received guidelines

within group settings had better calibration accuracy and higher exam scores than their peers

who were exposed only to one or neither of the interventions. In addition, DiGiacomo and Chen

(2016) used an intervention on sixth and seventh grade students in an experimental study that

targeted calibration practices across all three phases of Zimmerman's SRL model. They found

that students in the treatment group had significantly higher math performance and calibration

accuracy. Finally, Gutierrez and Schraw (2015) incorporated all three SRL phases in their

experimental study on undergraduate students. The researchers found significant effects for the

strategy training on performance, confidence, and calibration accuracy, and incentives further

improved performance and calibration accuracy.

From a theoretical perspective, SI sessions also target all three phases of Zimmerman's

(2002) SRL model. This is done by providing students with opening and closing session

activities that engage the forethought and self-reflection phases, respectively, and a primary

session activity that mirrors Zimmerman's performance phase. Since the SI model also aligns

with the three SRL phases, it was hypothesized that there would be a positive correlation

between SI participation and final calibration accuracy.

Furthermore, a variety of calibration research studies have indicated that high-achieving

students tend to be more accurate in their predictions than low-achieving students are (e.g., Bol

& Hacker, 2001; Flannelly, 2001; Nietfeld, et al., 2006; Shaughnessy, 1979). In addition, many

SI studies have found that students who attend SI tend to perform better in the course than their

peers (e.g., Grimm & Perez, 2017; Rabitoy et al., 2015). When coupling the findings of

calibration studies and SI research, it seemed likely that those who participate in SI would

perform better and have better final calibration accuracy than those who did not participate.

This study's path model results did not support this hypothesis. The lack of a relationship

between SI attendance and final calibration may indicate that the inherent connection between

the design of SI session plans and Zimmerman's SRL model is not strong enough to influence

students' calibration accuracy. Instead, SI leaders may need to incorporate intervention

strategies utilized by calibration researchers (e.g., reflection using group guidelines; Bol et al.,

2012) within their sessions to influence students' final calibration. This unanticipated result also

could be due to the previous findings in calibration research that people's confidence judgments

tend to remain stable over time (Hacker, et al., 2000; Nietfeld et al., 2006).

SI attendance does not influence final self-efficacy. The results of this study also

indicated a statistically non-significant relationship between SI attendance and final self-efficacy.

This supported some of the findings in the SI literature (Fisher, 1997; Garcia, 2006; Grier, 2004;

Visor et al., 1992; Watters & Ginns, 1997); however, it also contradicted Hizer's (2010) study

results.

In addition to examining participants' beginning self-efficacy and SI attendance, the study previously referenced by Hizer (2010) also examined self-efficacy at the end of the term. Results showed that, while students in the SI participation group had initially higher levels of anxiety, their anxiety decreased over the semester, while non-participants' anxiety levels increased. In addition, the researcher 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. The results of this study indicated that SI participation had a modest positive impact on self-efficacy for students in science courses who attended sessions regularly.

Conversely, a variety of SI studies found that SI attendance had no significant impact on students' final self-efficacy. In their study on early childhood major college students enrolled in a first-year foundational science course at an Australian university, Watters and Ginns (1997) found no significant differences in self-efficacy among students who attended and those who did not attend SI. However, the researchers did discover in their longitudinal study that students in 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. This means that students from the present study who frequently attended SI may have seen an increase in self-efficacy in their second semester of the general biology course.

Similarly, Fisher (1997) administered the MSLQ to undergraduate students in psychology courses, and results revealed no significant differences between the SI treatment and control groups on the self-efficacy scale. Grier (2004) also looked at the relationship between SI and

self-efficacy for a small sample of students in a grant-funded program, and analyses revealed no significant differences in self-efficacy among the four groups of (a) non-participants, (b) fall-only participants, (c) spring-only participants, and (d) both fall and spring participants. Lastly, Garcia (2006) employed a quasi-experimental study in which undergraduate anatomy and physiology students in existing courses were assigned to mandatory SI treatment and control groups. Using the Study Behaviors Inventory, Garcia found no statistically significant differences between the groups on the academic self-esteem scale.

To further add to the uncertainty of the potential influence of SI attendance on self-efficacy, Visor, Johnson, and Cole (1992), who published the first study to examine motivational factors as they relate to SI, found that SI participants saw a decrease in self-efficacy scores. While the results were not statistically significant, the researchers hypothesized that the decrease in SI participants' self-efficacy was because SI attendees better understood the rigor of the course and could reevaluate and adjust expectations of their ability. Conversely, Visor et al. speculated that nonparticipants "remained blissfully ignorant of what it takes to succeed" (p. 17). In other words, the authors connected an increase in students' calibration accuracy to a decrease in their self-efficacy. The findings from this study initially inspired the addition of calibration as a variable to the path model.

In summary, the existing literature made the effect of SI attendance on final self-efficacy unclear. Hizer (2010) found a positive relationship between SI attendance and self-efficacy, and Watters and Ginns (1997) found that high rates of SI attendance had delayed positive effects on students' self-efficacy during the second semester of a course sequence. However, three other studies found no significant effects of SI attendance on self-efficacy (Fisher, 1997; Garcia, 2006;

Grier, 2004). Finally, Visor et al. (1992) found a statistically non-significant decrease in SI

attendees' self-efficacy due to potential increases in calibration accuracy.

However, the present study's methodology did attempt to account for the shortcomings of

some of these studies. First, the self-efficacy scale used was reliable, valid, and related directly

to the general biology course. This was to account for the limitation of SI and self-efficacy

research studies that used instruments that were not task- or domain-specific to measure

students' self-efficacy (Fisher, 1997; Grier, 2004), which may have weakened the previous

studies' results (Pajares, 1996). In addition, unlike Fisher (1997), both a pre- and post-test were

administered to participants to control for potential differences in self-efficacy among SI and

non-SI participants at the beginning of the semester. The study also accounted for the limitation

in how different authors defined the SI group in varying ways (e.g., Visor et al., 1997 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) by using SI attendance as

a continuous predictor variable. Based on this study's methodology, it was anticipated that there

would either be a statistically significant positive relationship between SI attendance and final

self-efficacy or a negative relationship between SI attendance and self-efficacy by way of an

indirect influence of final calibration on final self-efficacy.

In the end, the path model indicated that SI attendance did not predict students' final self-

efficacy at the end of the semester, supporting the findings of Fisher (1997), Garcia (2006), Grier

(2004), and Watters and Ginns (1997) and differing from Hizer's (2010) results. Furthermore, SI

attendance did not predict final calibration, and final calibration and final self-efficacy were not

correlated. Thus, Visor et al.'s (1992) hypothesis that SI attendance increased calibration

accuracy thereby decreasing students' self-efficacy was not supported by this path model

analysis. However, the results of an independent samples t-test with the participant data did lend some credence to the claims of Visor and his colleagues. This additional analysis revealed statistically significant differences in SI attendance patterns between those who were over- and under-confident at the end of the biology course. This indicated that participants who attended SI more frequently were more likely to underpredict how well they would perform on the final exam, thus demonstrating a better understanding of the "severity of the challenge" (Visor et al., 1992, p. 17) in succeeding in the course.

SI attendance is correlated with improved final course grades. The results of the present study's path model indicated that SI attendance was a significant, positive predictor of participants' final course grades, supporting my hypothesis and previous findings by Rabitoy, Hoffman, and Person (2015) and Grimm and Perez (2017). Rabitoy et al. (2015) used linear multiple regression to discover 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. In addition, Grimm and Perez (2017) 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. The researchers found that SI attendance in both courses had a significant positive effect on course grades, even after controlling for prior achievement.

SI attendance also was predicted to have a significant, positive indirect effect on final course grade through improvements in final calibration accuracy and self-efficacy; however, this hypothesis was not supported. Previous research conducted by Nietfeld, Cao, and Osborne (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 on a small sample of undergraduate students in an educational psychology survey course, the researchers discovered a significant effect of average monitoring accuracy on self-efficacy. Since SI sessions encourages students' constant monitoring of their knowledge via activities that follow the forethought, performance, and self-reflection phases of Zimmerman's SRL model, it was anticipated that SI attendance would significantly improve students' calibration, performance, and self-efficacy. While participants who attended SI saw significant increases in their final course grades, this was not due to the indirect effect of SI attendance on calibration accuracy and self-efficacy. Instead, other factors influenced by SI attendance must have contributed to participants' increased course grades.

The Influence of SAT, Final Calibration, and Final Self-Efficacy

The path model used to explore the three research questions revealed several other statistically significant and not significant relationships among the variables outside of the primary research questions and hypotheses. Specifically, total SAT score had a significant relationship with nearly every variable in the path model with the exception of beginning self-efficacy. In this section, I first explore this and other phenomena among the exogenous variables – total SAT, beginning self-efficacy, and beginning calibration – as they relate to one another and the endogenous variables of SI attendance, final self-efficacy, final calibration, and final course grade. After this, I explore the relationships among the endogenous variables, including the most notable finding that final calibration and self-efficacy were positive predictors of participants' final course grades.

Exogenous variables: SAT influences most variables and students' calibration and self-efficacy are stable. The path model's exogenous, or independent, variables offer a variety of observations that fall outside of the research questions. Most notably, total SAT was

correlated with all of the variables except for beginning self-efficacy. In addition, participants'

beginning calibration and self-efficacy predicted their final calibration and self-efficacy.

One of the common findings in calibration research is that high-achieving students tend

to be more accurate in their predictions on assessments than their lower-achieving counterparts

are (e.g., Bol & Hacker, 2001; Flannelly, 2001; Nietfeld, et al., 2006; Shaughnessy, 1979).

Using total SAT score as a way of distinguishing between high and low achievers, it is not

surprising that the beginning calibration variable was correlated with total SAT score. In other

words, students with higher SAT scores (or high-achieving students) were more likely to predict

accurately their performance on their first exam.

Participants' beginning self-efficacy also was correlated with beginning calibration. This

finding is supported by Chen's (2003) observation that there is a positive significant relationship

between individuals' calibration accuracy and self-efficacy. Chen came to this conclusion after

studying the calibration and self-efficacy beliefs of seventh grade math students with a focus on

whether their calibration was a significant feature of their self-efficacy beliefs. Using a path

analysis, the study results demonstrated a significant direct effect of students' calibration

accuracy on their math performance and an indirect effect of calibration accuracy on students'

math performance through their math self-efficacy judgments.

While beginning calibration was correlated with beginning self-efficacy and total SAT

score in the path model, beginning self-efficacy and total SAT were not correlated. This was

unexpected, since self-efficacy refers to students' personal beliefs about their ability to perform

at certain levels (Bandura, 1977; Schunk, 1991; Schunk & Pajares, 2005). Thus, this finding

indicates that performance on the SAT must not influence students' beliefs in their ability to

perform well in biology. This may be due to the SAT being comprised primarily of math and verbal components and not specifically addressing biological sciences.

Participants from this study with lower SAT scores were more likely to attend SI. A few research studies have examined SI attendance patterns related to standardized test scores. While Congos and Mack (2005) found no significant differences in SAT scores between students in the SI and non-SI groups, most other researchers found that SI participants had significantly lower SAT (Peterfreund, et al., 2008), ACT (Hensen & Shelley, 2003), and AAR (ACT Aptitude Rating; Moore & LeDee, 2006) scores than non-SI participants. Total SAT score was, conversely, a positive predictor of final course grade for this study's participants. This finding is intuitive, and it is indicative of why standardized test scores are frequently used as control variables for prior achievement in SI research (Congos & Mack, 2005; Hensen & Shelley, 2003; Moore & LeDee, 2006; Peterfreund et al., 2008).

Another finding among the exogenous variables was that beginning self-efficacy positively predicted final self-efficacy, though the relationship was not very strong ($\beta = .38$). Nietfeld et al. (2006) found similar results in their study on monitoring exercises and feedback on calibration and test performance during an undergraduate course. Within Nietfeld's intervention study, the path model revealed a correlation between initial and final self-efficacy that also was significant but relatively low ($r = .33$).

As with beginning and final self-efficacy, students' beginning calibration also predicted final calibration, though the relationship was relatively weak ($\beta = .28$). Various calibration studies have found that people's confidence judgments typically remain consistent over time, regardless of their performance (e.g., Hacker et al., 2000; Nietfeld et al., 2006). Thus, it is not surprising that beginning calibration predicted final calibration, though one may have expected

the relationship to be stronger. In addition, once again, several calibration researchers have

observed that students with accurate calibration were more likely to be academically successful

(e.g., Bol & Hacker, 2001; Flannelly, 2001; Nietfeld, et al., 2006; Shaughnessy, 1979). This is

likely why beginning calibration had such a positive, significant impact on final course grade.

Finally, beginning calibration accuracy also was a positive predictor of final self-efficacy. While

this makes sense when considering the path model correlations between beginning calibration

and beginning self-efficacy as well as beginning and final calibration, the strength of this

relationship in the path model was unexpected due to the lack of previous research exploring the

relationship these two constructs.

       Endogenous variables: Final calibration and self-efficacy predict improved final

course grades. The path model results indicated that final calibration and self-efficacy were

both positive predictors of final course grades. Again, calibration researchers have found

previously that 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, et al., 2006; Shaughnessy, 1979). These findings were further supported by an

independent samples *t*-test using participant data. This analysis revealed that there were

statistically significant differences between those who under- and over-predicted their final exam

scores with higher average final course grades for students who were underconfident than those

who were overconfident. Similarly, self-efficacy research has demonstrated that college students

with high self-efficacy tend to have positive academic performances (Bandura et al., 1996;

Schunk, 2012).

In addition, while an examination of the exogenous variables revealed a correlation between beginning calibration and self-efficacy, there was not a significant relationship between final calibration and self-efficacy. This differs from a previous finding by Chen (2003) of a significant positive relationship between calibration accuracy and self-efficacy. This lack of a relationship between final calibration and self-efficacy was unexpected due to the significant relationships between beginning and final calibration, beginning calibration and self-efficacy, and beginning calibration and final self-efficacy.

Limitations

There are limitations to this study that should be considered when examining the results and their implications. First, as with most human subjects 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, total SAT scores were used in the SEM model.

Another threat to internal validity is social desirability, since the study uses self-report measures. I mitigated 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 was assured to participants during in-class announcements and via the electronic notification letter.

Students had the option to attend SI sessions led by two different SI leaders, which is another threat to internal validity. Fidelity was 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 weekly reviews of their session plans and session observations throughout the semester to ensure they are appropriately implementing

the SI model (see Appendices A and B for a sample session plan and SI observation record from the semester during which the study took place). While two SI leaders supported the students, a strength of this study is that one course instructor taught all the students.

A final potential threat to internal validity was attrition of study participants. I attempted 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. Overall, 320 students (66% of the class population) participated in this study, exceeding the 140-280 participants recommended using guidelines by Hancock and Mueller (2010) and Kline (2016). A related problem was low SI attendance, which can weaken the path model results. To combat this challenge, SI leaders made periodic in-class announcements and sent weekly reminders to students with session information. The course instructor also encouraged students to participate in SI. Overall, only 93 of the study participants (29%) attended SI at least once; thus, robust maximum likelihood estimation was used in Mplus (v 7.3) to address the non-normal distribution of the SI attendance data (Byrne, 2012).

The one-course, single-institution design of this study also threatens its external validity. Readers are cautioned on the generalizability of the study results to different contexts, and institutional context and detailed demographic information for study participants is provided for this reason. Further studies are encouraged to duplicate the procedures of this research to build external validity over time.

Implications for Further Research

Despite the internal and external limitations of this study, the findings provide several implications for future research. Specifically, I offer four areas of recommended additional research, including (a) replication of the current study, (b) further examination of additional SRL

variables that may affect final course grade indirectly because of SI attendance, (c) intervention studies related to SRL constructs and SI leader training, and (d) additional ways of approaching similar studies.

Replication of Current Study

This study has added to the existing literature in several unique ways, and it is important to continue the work begun with this research. First, this is the only known study to examine the potential relationships among calibration and SI, or any academic support intervention. The path model results demonstrated no significant relationship between students' calibration accuracy at the beginning of the semester and their decision to participate in SI, and there was no statistically significant relationship between SI attendance and participants' calibration accuracy at the end of the semester. However, additional analyses revealed that students who attended SI more frequently were more likely to underpredict how well they would perform on the final exam, while those who attended SI less often had more inflated views of their final exam scores. Additional research studies on SI and calibration, or more broadly on university academic support services and calibration, are needed. This can help expand our knowledge of students' metacognitive prediction abilities and their participation in voluntary academic support programs.

A second way in which the present study contributed to the literature was in the use of SI attendance as a continuous predictor variable, which has been done in a limited number of research studies (Grimm & Perez, 2017; Rabitoy et al., 2015). Most SI research uses SI attendance as a categorical variable in which students are divided into two or more groups based on SI attendance frequency (e.g., Blanc et al., 1983; Bruno et al., 2016; Terrion & Daoust, 2012). However, this artificial creation of categories may arbitrarily define the number of SI sessions

students must attend to experience changes in their academic performance. Cohen (1983) recommends using linear regression models to improve our understanding of the relationship between two variables, such as SI attendance and academic achievement. This recommendation demonstrates the need for additional multivariate studies that use SI attendance as continuous predictor.

A final way in which this study uniquely contributed to the existing literature was through its examination of the interaction between self-efficacy, calibration, and SI, as this was the first known study to include all three variables in an analysis. In addition, the path model findings added to the limited body of research that has explored the interaction between self-efficacy and calibration (Chen, 2003; Hong et al., 2014; Nietfeld et al., 2006), which are related but distinctive SRL concepts. In some ways, this research confirmed previous findings, including Chen's (2003) observation that there is a positive significant relationship between self-efficacy and calibration, which was true of students at the beginning of the semester. However, the lack of a statistically significant relationship between self-efficacy and calibration at the end of the term diverged from Chen's (2003) finding that students' beliefs are likely to remain stable over time regardless of actual performance. This finding also diverged from the research conducted by Nietfeld and his colleagues (2006) that showed that average calibration accuracy was a significant positive predictor of self-efficacy throughout a college term. Additional research is required to continue exploring how self-efficacy, a motivational construct, is related to calibration accuracy, a metacognitive factor.

Further Research on Other SRL Factors Influenced by SI

The path model results also indicate a need for more research on SRL variables that may affect final course grade indirectly through SI attendance. Based on the design of SI sessions

and previous research, I deduced that SI could positively influence participants' calibration accuracy and self-efficacy, which could better help explain the correlation between SI attendance and final course grade. However, the results of the adjusted path model indicated that, while SI attendance, final calibration, and final self-efficacy were all significant predictors of an increased final course grade, there was no significant relationship between SI attendance and final calibration or self-efficacy.

Using Zimmerman's model for self-regulated learning (2000, 2002), there are several other elements that one may consider when looking at the potential indirect influence of SI on final course grade through SRL. First, there are other self-motivation beliefs outside of self-efficacy that are present in the forethought phase of Zimmerman's model, including outcome expectations, intrinsic interest/value, and learning versus achievement goal orientations. Mack (2007) used the MSLQ to discover that chemistry students who frequently attended SI (8+ sessions) had a positive correlation with the motivation scale, which included intrinsic motivation, extrinsic motivation, goal orientation, task value, control of learning beliefs, self-efficacy, and test anxiety. In addition, McGee (2005) found statistically significant correlations between student participation in SI on extrinsic motivation, self-efficacy, and control beliefs using the MSLQ. This may point to other motivational constructs being more affected by SI participation than self-efficacy.

In addition to examining more closely the self-motivation beliefs present in the forethought phase, the performance phase of Zimmerman's (2000, 2002) model is another area in which additional SI research may be focused. Specifically, McGee (2005) also found statistically significant correlations between SI attendance and the organization, effort regulation, peer learning, and help-seeking scales on the MSLQ. In addition, Fisher's (1997) study revealed

statistically significant differences between SI treatment and control groups on the peer-learning

and help-seeking scales of the MSLQ, which is intuitive when one considers that SI is a

voluntary, peer-led academic support program.

Finally, additional research using a regression or ANOVA analysis could examine the

influence of SI attendance on the final phase of Zimmerman's SRL model: self-reflection. This

phase involves learners engaging in self-judgments, including comparing one's performance to a

perceived standard and attributing successes or failures to internal or external factors. The self-

reflection phase also consists of self-reactions, including people's felt satisfaction or

dissatisfaction with their performance and adaptive or defensive inferences from their

performances. There has not been much research conducted on the influence of SI attendance on

students' self-reflection abilities, which makes this another potential area for future studies.

Intervention Studies on SRL and SI Leader Training

Intervention studies related to SRL constructs and SI leader training provide a third major

area for future research. Much of the literature points to theoretical connections between the SI

model and theories of learning like SRL; however, many of the resources provided for training

SI leaders does not help them directly understand how their sessions can help improve students'

SRL abilities and why this is important. Specifically, with special training, SI leaders may be

able to focus their introduction activities on task analysis and improving students' self-

motivation beliefs. Then, they could direct their primary session strategies to helping students

develop self-control and self-observation skills. Lastly, closing session activities could be better

designed to help students self-reflect on their learning behaviors during and outside of the SI

session (Zimmerman, 2000, 2002). It would be interesting to study the differences between how

SI leaders facilitate their SI sessions before and after going through such a targeted training, as well as the potential effects on their SI participants' SRL abilities.

Similarly, SI leader training in self-efficacy theory and research could help them understand the four primary sources of influence on students' self-efficacy: mastery experiences, vicarious experiences, social persuasions, and emotional and physiological states (Bandura, 1977; Usher, 2009). Since mastery, or performance-based information, is the most powerful influencer (Schunk, 1991), SI leaders could be shown how to scaffold student learning through hands-on activities. While this may be covered implicitly in SI leader training, an intervention study that compares explicit versus non-explicit SI leader training in these areas and the differences in student self-efficacy and SI session activities may be worthy of further study.

Finally, an intervention study could be conducted in which SI leaders are trained to put specific calibration research strategies into practice within their SI sessions to see if there is any impact on participants' calibration accuracy and course performance. For example, Bol et al. (2012) discovered that participants in their study who were provided with guidelines in group settings had better calibration accuracy and higher exam scores than their peers who were not provided with guidelines and/or who studied in individual settings. Similarly, DiGiacomo and Chen (2016) found that students had significantly higher math performance and increased calibration accuracy after being exposed to a set of structured, guided questions with feedback and self-reflective worksheets. These types of calibration activities could be implemented in SI leader training and sessions to conduct additional research on their effect in an academic support setting.

Additional Approaches to Similar Studies

Finally, there are additional ways in which similar research could be conducted to build upon the work of the present study. This could include a longitudinal study, additional demographic factors, or a quasi-experimental study with randomized and control groups.

Watters and Ginns (1997) published a longitudinal study that examined the impact of SI on the self-efficacy of early childhood major college students enrolled in a two-semester foundational science course series. The researchers discovered no statistically significant differences in self-efficacy for students who participated in SI at the end of their first semester in the course. However, the authors administered the Science Teaching Self-Efficacy Belief Instrument to the students again at the end of their second semester and found that students who had attended more than 66% of the offered SI sessions during the first semester had significant increases in self-efficacy the following semester. This indicates that the effects of SI participation on students' self-efficacy may take time to develop. For this reason, a recommended area for future research would be replication of the current research study with an added longitudinal approach. The path model used for this study could be extended to include SI attendance and final calibration, self-efficacy, and course grades for students enrolled in the second semester of the General Biology course.

Another potential extension of the present study would be to include additional demographic factors in the path model. In the current study, gender and race/ethnicity were not used as path model variables because it was unclear if the sample size would be large enough to account for the additional variables (Hancock & Mueller, 2010; Kline, 2016; StataCorp LLC, 2018). However, 320 participants exceeded the minimum of 140-280 participants required to achieve reliable results, indicating that additional variables could have been included in the path

model. Other researchers have looked at the potential unique effects of SI attendance on student academic performance based on gender (Fayowski & MacMillan, 2008; Mack, 2007) and race/ethnicity (Mack, 2007) to identify if the support program affects students from different backgrounds in varying ways. In these studies, there were no statistically significant differences in the effects of SI attendance on academic performance based on gender (Fayowski & MacMillan, 2008; Mack, 2007) or race/ethnicity (Mack, 2007). A related alternative option for a research study would be to examine the impact of SI on students taking courses within different disciplines, such as chemistry, math, or history. For example, Mack (2007) separately analyzed students enrolled in chemistry and biology courses and found that frequent SI participants specifically in chemistry had statistically significantly higher levels of motivation at the end of the semester when compared with occasional and non-SI participants.

A final potential variation to extend the present research study could include the utilization of randomized and control groups to produce a quasi-experimental research study. This type of research design can offer more control, but it was not practical at the institution used for the present study, due to the longstanding SI support offered to all students enrolled in general biology. However, a quasi-experimental study may be more plausible at an institution in which SI is not an expected support structure for all students enrolled in a course. Garcia (2006) utilized a quasi-experimental research design in which students in existing anatomy and physiology courses were assigned to mandatory SI treatment or control groups that received chapter-specific web-based reviews. Similarly, Fisher (1997) examined students in three Psychology courses in which participants in only one of the course sections had access to SI, while the other courses served as control groups.

Implications for Practice

The results of this research study also provide implications for practice. In this section, I address the practical value of SI as an academic intervention program; training opportunities for SI leaders; and teaching interventions that may be employed by instructors of high-risk, introductory college-level courses.

Value of Supplemental Instruction for High-Risk Courses

First, while the study did not produce all the expected results from the hypothesized path model, it added to the growing body of literature that highlights the positive impact of SI on students' final course grades in high-risk courses (e.g., Arendale, 1997; Blanc et al., 1983; Grimm & Perez, 2017; Martin & Arendale, 1992; Rabitoy et al., 2015). Thus, while SI may not have influenced specific areas of students' SRL abilities, it still positively influenced their term GPA and ability to persist in a course that many students struggle to pass. Specifically, students were able to increase their course GPA by approximately half a letter grade by attending SI on a weekly basis. Therefore, SI remains a viable support program option for institutions exploring ways to help students pass high-risk courses that have high DFW rates, and SI programs should be continued in their present form to support students in these challenging courses.

Research-Based SI Leader Training Redesign to Target SRL and Self-Efficacy

SI leader training content is a second area of practice that may be influenced by this research study. The literature review provided detailed information on Zimmerman's theory of self-regulated learning (2000, 2002) and the clear practical connections between this theory and the design of SI leaders' session plans (Curators of the University of Missouri, 2011). The review of literature also addressed how the SI model has the potential to affect positively the four

primary influencers of students' self-efficacy: mastery experiences, vicarious experiences, social

persuasions, and emotional and physiological states (Bandura, 1977; Usher, 2009).

While most SI leader training sessions emphasize how to plan session activities and

utilize facilitation skills (e.g., redirection, wait/think time, and checking for understanding), there

often is not a clear connection to these practices and research-supported theory. If SI leader

trainings were redesigned to help the leaders understand the SRL and self-efficacy theories that

inform their session activities and facilitation strategies, they may be more mindful of how they

implement these practices.

For example, sharing Zimmerman's SRL theory could help SI leaders better understand

the value of effectively managing their session time to allow for a beginning and closing activity

during each session. Similarly, training on the four sources that influence self-efficacy could

encourage SI leaders to do more intentional modeling and scaffolding of their session activities

to help students build up their vicarious and mastery experiences.

Teaching Interventions for Instructional Faculty

Finally, the path model revealed statistically significant positive relationships between

students' final self-efficacy and calibration accuracy and their final course grades. Thus, a

valuable implication for practice is to inform instructors of high-risk courses, like general

biology, of research-based teaching practices they could implement to improve students' self-

efficacy and calibration accuracy.

In the review of the literature, several recommended instructional strategies were shared

to help increase students' self-efficacy. For example, students' mastery experiences can be built-

up by providing them with challenging, meaningful tasks that they are capable of mastering and

by offering plenty of support, encouragement, and autonomy throughout the learning process

(Pajares, 2002). In addition, faculty can encourage an increase in students' self-efficacy by providing feedback for early successes, tangibly rewarding successes, and explicitly pointing out to students how prior learning in the course has prepared them for new content (Schunk, 1991, 2012). Faculty also can practice modeling for their students, including emphasizing specific SRL and learning strategies that can help them succeed in the course and demonstrating that it is okay to make mistakes (Pajares, 2002; Schunk, 1991). A final instructional tool that course instructors can use to promote self-efficacy is to help students set learning goals that are short-term, specific, and start off easy before becoming progressively more challenging (Pajares, 2002; Schunk, 1991).

In a similar manner, the calibration research offers several ideas for teaching strategies that have had a positive effect on students' calibration accuracy and academic performance. For example, Bol and her colleagues (2012) found that high school biology students benefited positively when provided with guidelines and when working through these guidelines within group settings. Similarly, DiGiacomo and Chen (2016) provided sixth and seventh graders with structured, guided questions that helped them review course material and calibrate their performance. In addition, students received feedback and used self-reflective worksheets. Finally, Gutierrez and Schraw (2015) were able to improve students' calibration accuracy and academic performance by incorporating specific cognitive strategies of instruction related to calibration, incentives for high performance to increase motivation, and self-reflection in the form of confidence judgments after completing items on an assessment. Faculty teaching students in challenging courses could review calibration and self-efficacy research for ways of incorporating these teaching strategies into their classrooms. In addition, institutions with

centers for faculty development or teaching and learning could incorporate these research-based strategies into faculty workshops.

Conclusion

In this final chapter, I summarized the study's major findings on Supplemental Instruction, self-efficacy, and calibration, including the results of the three research questions and other observations from the adjusted path model. After providing a summary, these findings were contextualized within the literature on Supplemental Instruction, self-regulated learning, self-efficacy, calibration, and help seeking, again addressing the three research questions and other outcomes among the path model variables.

The results of this study provided several interesting observations. In addressing the first research question, I found that students with low self-efficacy were more likely to attend SI during the semester; however, students' calibration accuracy at the beginning of the term had no influence on their SI attendance. Results related to the second and third research questions revealed that, while SI positively influenced students' final course grades, this relationship could not be attributed to any effects of SI attendance on final self-efficacy or calibration accuracy. Outside of the research questions, total SAT score had a significant relationship with the other variables in the path model except for beginning self-efficacy. Self-efficacy beliefs and calibration accuracy also remained stable for students from the beginning to the end of the semester. Finally, students' calibration accuracy and self-efficacy at the end of the semester were both positive predictors of final course grade.

After sharing the results of this study within the context of the literature, I outlined several potential threats to internal validity, including self-selection bias, social desirability issues, potential fidelity challenges, and attrition of study participants. In addition, a threat to

external validity via a one-course, single institutional study design was presented. I also shared

several implications for further research, including replication of the current study, further

examination of additional SRL variables that may indirectly affect final course grade,

intervention studies related to SRL constructs and SI leader training, and additional ways of

approaching similar research. Finally, three implications for practice were offered, including the

value of SI as an academic support program for high-risk courses, suggestions for training SI

leaders to target SRL and self-efficacy, and teaching interventions for instructional faculty.

While not all the path model results were expected, this study contributes new empirical

and practical viewpoints to the literature on Supplemental Instruction, calibration, and self-

efficacy. In addition, several areas for suggested further research and implications for practice

may be of interest to educational psychologists, higher educational professionals, and college and

university faculty and administrators.

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

Congos, D., & Mack, A. (2005). Supplemental Instruction's impact in two freshman chemistry

classes: Research, modes of operation, and anecdotes. *Research & Teaching in

Developmental Education, 21*(2), 43-64.

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.

Grubbs, F. E. (1969). Procedures for detecting outlying observations in samples. Technometrics, 11, 1–21. doi: 10.1080/00401706.1969.10490657

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. (2019). Calibration and self-regulated learning: Making the connections. In J. Dunlosky & K. Rawson (Eds.). *Cambridge handbook of 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.

Hensen, K. A., & Shelley, M. C. (2003). The impact of Supplemental Instruction: Results from a

large, public, Midwestern university. *Journal of College Student Development, 44*, 250–

259. doi:10.1353/csd.2003.0015

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.

Keith, T. Z. (2015). *Multiple regression and beyond: An introduction to multiple regression and*

        structural equation modeling (2nd ed.). New York, NY: Routledge, an imprint of Taylor

        & Francis.

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

Moore, R., & LeDee, O. (2006). Supplemental Instruction and the performance of developmental

education students in an introductory biology course. *Journal of College Reading and

Learning, 36*(2), 9–20.

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., Brooks, W., White, A., Richmond, E., Cromley, J. G., Kaplan, A., Dai, T., … Balsai,

M. (2018, April). Do perceived costs affect achievement in an undergraduate biology

course? It depends on self-efficacy. Poster session presented at the meeting of American

Educational Research Association, New York City, NY.

Peterfreund, A. R., Rath, K. A., Xenos, S. P., & Bayliss, F. (2008). The impact of Supplemental

Instruction on students in STEM courses: Results from San Francisco State University.

Journal of College Student Retention: Research, Theory and Practice, 9, 487–503.

doi:10.2190/CS.9.4.e

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

Rask, K. (2010). Attrition in STEM fields at a liberal arts college: The importance of grades and pre-collegiate preferences. Economics of education Review, 29(6), 892-900. doi: 10.1016/j.econedurev.2010.06.013

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.

Serra, M. J., & DeMarree, K. G. (2016). Unskilled and unaware in the classroom: College

    students' desired grades predict their biased grade predictions. Memory & Cognition, 44,

    1127-1137. doi: 10.3758/s13421-016-0624-9

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

SAMPLE SI LEADER SESSION PLAN

Date of Planning: 9/18/18

PASS SESSION PLANNING

PASS Session Date: 9/20/18

Course Instructor: ███████

Course Name: BIOL121N

PASS Leader: ███████

Objective: What does this group most need to accomplish in this session?

Identify the active site, substrates, and products formed with enzymes; understand that enzymes are biologically necessary for life and how they work.

	Content to cover:	Process to use:
Introduction	Sign in See how first test went (if they've taken it) Intro to enzymes	KWL: After checking to see how students felt about the first test, I will draw a KWL chart on the board and explain what each column stands for (know, want to know, learned). I will ask each student to write/add at least 2 things in the K column and 1 thing in the W column (depending on attendance). This should take between 5-10 mins.
Main Activity(ies)		Labelling/drawing diagrams: On the board, I will have a blank diagram that shows the catalytic enzyme process (labels covered up), and the students will have to label and important pieces and explain what each piece does in function (~10 mins).

	Enzymes	Board race: I will have the mini dry erase boards and extra markers. If there are enough students in attendance, I will break them into groups. If not, then partners will work too. I will have about 15 practice questions and/or definitions prepared. Students will have to quickly write the answer or word on the board and hold it in the air (~35 mins).
Closing Activity	Wrap up enzymes	Bookend with L portion of KWL chart: Students will have to share at least one thing that they now know about enzymes at the end of the session that they didn't remember before walking in. They will add this to the L column of the KWL chart that we began in the session (~5 mins).

AFTER SESSION REFLECTION

1. **Did the students grasp the material well? Do you feel as if the topic needs to be covered again? Explain.**
   The student who attended my session was not particularly confident in her knowledge at first, but was able to tell me at the end of the session about the things she learned. I think that this topic should be covered again just so that I can reach more students -- I may dedicate part of a session to enzymes but I don't think that I would spend another whole hour on them.

2. **How did students react to the activity? Do you think you will use the same activity again? Explain.**
   Since only one student showed up, the activities I had planned were not as fun as I was hoping they'd be. I think the labeling was very useful, and I will definitely do this activity again. I think that the board race would be useful if there were at least a handful of students in attendance.

3. **What study tips did you share?**
   Watch the simulations/animations in the homework assignments/note slides because they demonstrate the more difficult processes in a simpler way. Create visual diagrams and concept maps to help yourself when studying.

4. **Other session comments/thoughts (I.E. attendance):**
   Only one student showed up. Today was the day I was evaluated -- I was hoping for higher attendance.

APPENDIX B

SAMPLE COMPLETED SI LEADER OBSERVATION RECORD

PASS Session Observation Record – Fall 2018

Leader Name:	█████████
Course/Section:	BIOL ███
Date/Time/Location:	9/20/18
Number Attending:	1
Observer:	█████████

Administrative Tasks:
(e.g., session starting on time, PASS sign, participation log, necessary supplies)

Sign was posted early. Easy access to sign in.

Opening the Session:
(e.g., setting the agenda, asking for student questions/input, opening activity)

Started by setting the agenda. Moved into a KWL chart. She helped the student build self confidence.

Processing Activities:
(e.g., group collaboration techniques, recall & review, organizational & visuals, study techniques, problem solving strategies)

Labeling/ drawing diagrams on the board. The student did well with this activity. Coming into the activity she was very unsure of herself and the answers. After this activity the student had a better understanding of the different parts of enzymes and how they worked.
She also had planned a board race but since there wasn't enough students she modified it so that her and the student went over the questions together. She used many facilitation skills.

Facilitation Strategies:
(e.g., redirection, questioning techniques, wait/think time Types 1&2, checking for understanding, participation strategies)

She did well on wait time type 1. She could use practice on wait time type 2. She did well with redirection for only having 1 student. She did well on checking for understanding. She could use practice on slowing her speed of speech down.

Session Atmosphere:
(e.g., rapport with students, managing of student emotions, students comfortable participating/asking questions, students interacting with one another)

The atmosphere was very relaxed. She had a good rapport with the student and made her feel comfortable. The student showed a lack of self confidence and she helped build her confidence.

Closing the Session:
(e.g., management of time, reviewing what was discussed, allowing time for questions, previewing what's coming up in class, marketing future sessions)

She finished the L for the KWL chart. This activity was good for checking for understanding.

Session Strengths:

Sitting with the student. Helping to build self confidence. Extremely personable.

Thoughts for Future Sessions:

Looking up the information together when there is only one student. Try not to talk too fast.

APPENDIX C

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 D

COURSE INSTRUCTOR DATA REQUEST LETTER

Dear ███████

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 ███████ 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 not to 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 E

OFFICE OF INSTITUTIONAL ASSESSMENT DATA REQUEST LETTER

Dear ███████

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 ███████ 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 F

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 ▮▮▮▮▮▮ 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

VITA

JENNIFER L. GRIMM

Student Success Center, Old Dominion University ▪ (757) 683-7651 ▪ jgrimm@odu.edu

EDUCATION

Doctorate of Philosophy   **OLD DOMINION UNIVERSITY**
Program: Higher Education   *August 2019; Norfolk, Virginia*

Master of Education   OHIO UNIVERSITY
Program: College Student Personnel   *June 2011; Athens, Ohio*

Bachelor of Business Administration   **OHIO UNIVERSITY**
Double Major: Marketing/Human Resource Management   *June 2009; Athens, Ohio*

PROFESSIONAL EXPERIENCE

Director of Academic Initiatives   *May 2019 to Present*
Center for High Impact Practices, Old Dominion University *Norfolk, Virginia*

Director of the Peer Educator Program (PEP)   *September 2014 to May 2019*
Center for High Impact Practices, Old Dominion University *Norfolk, Virginia*

Supplemental Instruction (SI) Coordinator   *August 2012 to August 2014*
Academic Resources, Carroll University   *Waukesha, Wisconsin*

Hail Hall Residence Director   *June 2011 to May 2012*
Residence Life, Belmont University   *Nashville, Tennessee*

Graduate Assistant *September 2009 to June 2011*
Office of the Dean of Students, Ohio University   *Athens, Ohio*

PUBLICATIONS

Grimm, J. (2015). Overcoming Ignorance: Coming to an Understanding of How White Privilege Has Impeded Efforts for Educational Desegregation. *ABC-CLIO's Enduring Questions Series, The African American Experience*.

Smith, K. J., **Grimm, J.**, Lombard, A. E., Wolfe, B. (2012). Cyberbullying: It doesn't stop after high school graduation. In L. A. Wankel & C. Wankel (Eds.), *Misbehavior online in higher education: Cutting-edge technologies in higher education*. Bingley, UK: Emerald Group Publishing Limited.

PROFESSIONAL PRESENTATIONS

Grimm, J., Mize, M., Forbes, B. (2019). Leading the Way: Crossing Divisional Boundaries to Prepare Tomorrow's LeADERS. Paper presented at the American Council on Education's Virginia Network Conference, Harrisonburg, VA.

Grimm, J., Moser, L. (2018). Leveraging Online PASS Leader Training to Promote Integrative and Collaborative Learning. Paper presented at the SI International Conference, Seattle, WA.

Grimm, J., Perez, T. (2017). A Longitudinal Study of Supplemental Instruction's Impact on Anatomy and Physiology. Poster presented at the American Educational Research Association (AERA) Annual Meeting, San Antonio, TX.

Grimm, J., Reid, T. L. (2016). Aligning Tutoring Center Objectives, Practice, and Assessment: From Theory to Practice. Paper presented at the College Reading and Learning Association Annual Conference, Louisville, KY.

Grimm, J. (2014). Save the Trees: Strategic Use of Digital Resources for the 21st Century SI Program. Paper presented at the SI International Conference, Chicago, IL.

Grimm, J. (2014). Supporting Anatomy and Physiology Students in Transition. Paper presented at the Wisconsin Learning Assistance Network Conference, De Pere, WI.

Lombard, A. E., **Grimm, J.** (2013). Cyberbullying in College. Online webinar presented for the Association of Student Conduct Administration.

Lombard, A. E., **Grimm, J.** (2012). Cyberbullying: It Doesn't Stop After High School Graduation. Paper presented at the ACPA: College Student Educators International Annual Convention, Louisville, KY.

HONORS AND AWARDS

Selected for the "Scholarship Award" among the Old Dominion University Higher Education Ph.D. Program graduates for my academic writing and contributions (May 2019)

Selected as the "Outstanding Student in College Student Personnel" among the 25 graduates of the Ohio University College Student Personnel M.Ed. Program (May 2011)

Selected as the "Outstanding Student in Human Resources" from all graduating seniors within the Ohio University Human Resource Management major (May 2009)