TRACKING AND UNDERSTANDING AT-RISK COLLEGE STUDENTS' SELF-EFFICACY
IN A TREATMENT MATHEMATICS COURSE

By

Lauren Renee Cavner Williams

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Mathematics Education – Doctor of Philosophy

2020

ABSTRACT

TRACKING AND UNDERSTANDING AT-RISK COLLEGE STUDENTS' SELF-EFFICACY
IN A TREATMENT MATHEMATICS COURSE

By

Lauren Renee Cavner Williams

Students enrolled in developmental college courses generally have lower levels of self-efficacy than their counterparts, and this can negatively affect their course performance, motivation, and ability to graduate. The current study focuses on the effects of a treatment course in developmental algebra on students' self-efficacy in mathematics. This course is part of a program that assists students who place into developmental mathematics, English, and reading courses with additional supports and saving students time and money. The perceived effects of the treatment developmental algebra course on students' self-efficacy in mathematics was measured with a standard instrument (the Mathematics Self-Efficacy Scale [MSES]), student journals, student interviews, student grades on course assessments, and the instructor's journal. The MSES was also administered to similar non-treatment students to offer further comparison. Data analysis determined that changes in self-efficacy occurred and were measurable, more so for the treatment group than the control group. Additional analysis points to the reasons that treatment students may attribute to these observed changes in self-efficacy which include opportunities of success and persistence through struggle, a course instructor who is inviting of questions and willing to help students, and peer support from the learning community. Further, data analysis illustrated the level of agreement between the results of the five data sources. Results of this analysis indicate that agreement generally existed between the data sources, but

inclusion of journals potentially provides details about the student experience that might have

otherwise been lost.

ACKNOWLEDGEMENTS

Thank you very much to my dissertation committee Dr. Jack Smith, Dr. Monica Smith Karunakaran, Dr. Lisa Linnenbrink-Garcia, and Dr. Vince Melfi for your support and thoughtful feedback throughout my dissertation. A special thank you to my advisors, Jack and Monica, for our frequent meetings, support, and guidance as I developed my proposal, collected and analyzed my data, and wrote this dissertation. I cannot express how much your support and encouragement means to me!

Thank you very much to my two awesome student data collectors, Andrea and Chris. You guys were so dedicated and supportive throughout the entire data collection phase. I could not have done this without your help!

A big thank you to Dr. Stacy Priniski for all of your help and support with SPSS. Your willingness to meet with me, answer questions, and provide feedback was so generous and I cannot thank you enough!

To my fellow graduate student Valentin Kuechle, thank you so much for helping me establish inter-rater reliability. Your dedication to helping me and thoughtful discussions greatly benefited my work and I hope to return the favor in the future!

Thank you very much to my C4 colleagues: Dr. Kirk Weller, Dr. Heather Pavletic, and Prof. Kristin Conley. Your support, encouragement, and patience working with me is most appreciated!

Also, this study would not have been possible without the participation from all of the students involved in my treatment and control groups. Thank you for participating!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Programs have been established to assist students who enroll in college but are not quite ready for the demands of college-level academic work and/or a more independent lifestyle. To help these students, often in their first year, colleges offer remedial courses, also known as developmental courses (Zientek et al, 2019). However, these students are statistically at greater risk of never completing their desired degree (Benken et al., 2015). Contributing factors to non-completion include greater time commitment to complete additional credits and additional costs (Benken et al., 2015). Failure rates of these courses tend to be particularly high, especially in developmental mathematics (Zientek et al., 2019). This can be linked to many factors including one's self-efficacy toward the subject.

Self-efficacy is the focus of the current study because it predicts many educational outcomes including academic achievement (Fong et al., 2015; Katz, 2015; Usher & Pajares, 2008; Yurt, 2014; Zientek et al., 2019), and more specifically GPA (Fong et al., 2015), college major choice (Usher & Pajares, 2008), and career choice (Bong & Skaalvik, 2003; Usher & Pajares, 2008). Self-efficacy has been seen as the "most useful self-construct because it predicts subsequent motivation and performance better than the other constructs, including self-concept" (Bong & Skaalvik, 2003, p. 29). To further highlight the construct's importance, students with higher self-efficacy have been shown to "work harder, persist longer, persevere in the face of adversity, have greater optimism and lower anxiety, and achieve more than low-efficacious students. They approach difficult tasks as challenges to be mastered rather than as threats to be avoided" (Katz, 2015, p. 43). Therefore, lower self-efficacy can be detrimental to student success, as it causes doubt in a student's mind about their ability to complete a task (Hall & Ponton, 2005). Alternatively, higher self-efficacy promotes confidence and students are more

likely to persevere when given a task, even if the level of difficulty is high (Hodges, 2005). A problem that plagues developmental mathematics classrooms is the lower levels of self-efficacy that enrolled students generally possess (Smith, 2010), especially when compared to their non-developmental counterparts (Hall & Ponton, 2005). This can negatively influence student performance, motivation, and perseverance.

The current study, which was designed and carried out as a quasi-experiment, focused on this problem by seeing what, if any, changes occurred in self-efficacy toward mathematics when college students took a special section of a developmental mathematics course. This course was an intervention that differed from the standard mathematics program, and the current study assessed its effects by comparing students in the "treatment" section to similar students in the standard mathematics course sequence. The "treatment" course was designed to give students experiences of success with mathematics, focus on the content and skills they typically struggle with, and provide them with support as they persist and/or progress through the course. Further, the "treatment" section differed from a regular section of the same course in terms of materials and content, structure, instructional practices, and assessment. These differences will be outlined later in this document.

The participants in the treatment condition were students in a cohort program at a Midwest university that taught special sections of developmental mathematics, English, and reading. Students enrolled in these developmental courses were considered at-risk (Benken et al., 2015). This cohort program, called the Cross-Curricular Career Community (C4) Scholar Program, focuses on retention of the at-risk student population at the university by instilling accountability and professionalism in the students. This program enrolls students into courses that are one level above their initial placement and provides extra supports that students normally

would not receive in a traditional mathematics course. These supports, described in Chapter 3, work to provide students with opportunities for success and help to persist through struggle. In addition, students save time and money because of the higher course placement.

I, the author and researcher in the study, taught the mathematics course in the C4 Scholar Program for three academic years (co-instructor for 2016-2017 and 2017-2018; sole instructor for 2018-2019) prior to undertaking the study. In my experience, students come into the developmental algebra course with feelings associated with negative self-efficacy for mathematics. During my time teaching and interacting with past cohorts of students, I thought that I saw positive changes in many students' self-efficacy towards mathematics occurred. But I had no evidence to support or question this teacher's hunch. This was and is what motivated the current study in order to understand more about the changes in students' self-efficacy and whether they were really happening. This study attempts to move beyond this teacher's hunch by using and developing different measures of self-efficacy in mathematics.

There were four research questions that guided the current study:

1. What, if any, changes in mathematics self-efficacy occur for at-risk college students, as measured with a standard instrument, when the instructional treatment is implemented?

2. How do these results compare to students of similar placement, but who do not receive the treatment, as measured with the same instrument?

3. According to students' journals and interviews from the treatment group, have changes in their level of self-efficacy occurred? And, what do they think are the reasons for these changes or absence of changes?

4. What consistencies and/or or inconsistencies exist between the results of each of the collected data sources: (a) the standard instrument (*MSES*), the students' perspectives

through (b) journals, (c) interviews, (d) students' course assessments, and (e) the instructor's perspective, which was motivated by documented instances of possible influence on participants' self-efficacy throughout the semester in daily instructor journals?

The first research question was constructed to reveal if the treatment course of the current study was effective in increasing the self-efficacy of at-risk first year college students, as measured by a standard instrument. The second research question offers a comparison to a similar group and their responses on the same standard instrument to see if the results of the first question could have been because of participants' enrollment in the treatment course. The third research question was designed similarly to the first research question, but results were measured with qualitative data collected methods. The fourth research question focuses on comparing the results from analysis of the five data sources to see if agreements or disagreements result.

To answer these questions, I, first, present a literature review in Chapter 2 which focuses on what is known about the at-risk first year college student population, developmental mathematics courses in college, self-efficacy, and interventions enacted to improve this group's self-efficacy toward the subject. Chapter 3 reviews the methods of data collection and data analysis that I used for the current study. Chapter 4 reveals the results I obtained from data collection and data analysis. Finally, Chapter 5 discusses the results, provides implications for future research, and discusses the limitations of the current study.

CHAPTER 2: LITERATURE REVIEW

This chapter will help to situate the current study within the field of mathematics education by synthesizing what is known about first year college students who are at-risk, the challenges they encounter, the importance of mathematics and developmental mathematics courses, and the importance of understanding these students' self-efficacy toward mathematics. The reader should notice a broad to narrow focus, beginning with the general first year college student and eventually narrowing to the self-efficacy of at-risk first year college students who take developmental mathematics courses, as it is the intention of this chapter.

Common Challenges for First Year College Students

The transition to college is a challenge for many students. Some of the more common challenges include the unpreparedness of incoming students, their use of ineffective study strategies, and issues with engagement in their courses. This section will discuss these challenges, because they are things instructors should be aware of when working with first year college students.

Many college instructors would agree that their first-year college students often come to college unprepared. This includes those at community colleges and those at four-year universities. Many researchers have also reported this (Alward, 2012; Benken et al, 2015; Brewer, 2009; Deshler & Fuller, 2016; Fong et al., 2015). Students are not academically ready for the challenges that college courses demand of them. Benken et al. (2015) indicates that even though "some students are successfully passing high school mathematics courses and are attending postsecondary institutions, they are clearly not ready for success within mathematics at a level comparable to when they leave high school" (p. 15). One way that colleges have responded to this issue is by creating developmental courses and programs to support students

who want to pursue a degree but who are not quite ready for mainstream coursework. What these developmental courses look like will be discussed later in the review.

Another challenge that first year college students encounter is the use of ineffective study strategies. Many students come to college claiming that they rarely or never studied in high school, or at least never did any work outside of the classroom. One strategy that many students could improve on is note taking. Cardetti et al. (2010), for example, focused on how students take notes and found that many are actually incomplete for their coursework. Notes are an important component in the college classroom and Cardetti et al. reported that "successful students included more of the main ideas in their notes than did less successful students" (2010, p. 80). To help instructors support their students, Cardetti et al. (2010) recommended that instructors create and hand out partial notes for use during instruction. This recommendation extends to mathematics courses. Another skill that these students could improve, and would help their preparedness, is "transferring mathematical ideas across disciplines" (Christensen & Thompson, 2012, p. 5) and outside of the classroom. A reason for this issue is often related to students not making necessary connections between concepts (Nimela et al., 2018).

First year college students also have issues with feeling engaged in their courses (Alward, 2012). According to Alward (2012), contributing factors include not feeling a personal connection to the college or university, distractions in their outside lives including relationships, medical and financial issues, and feeling overwhelmed with the workload and/or lifestyle that accompanies college. If students feel as though they do not belong or have support at school, are distracted by personal issues, or do not feel that they will succeed then it is likely that their engagement level will be negatively impacted. Factors like these can consume the energy that students may have otherwise invested toward engagement in their courses. In particular,

according to the US Department of Education (2018), 43% of full-time and 81% of part-time undergraduate students in 2018 were employed. Of full-time students who were employed, 17% worked 20-34 hours per week and 10% worked 35 or more hours per week. Of part-time students who were employed, 24% worked 20-34 hours per week and 47% worked 35 or more hours per week. Therefore, the percentage of students who work 20 or more hours per week is considerable and the need to work can extensively affect the academic performance of many students. In a study by Pike, Kuh, and Massa-McKinley (2008), the authors concluded that "students who work more than 20 hours per week have significantly lower grades than students who do not work," which is tied closely to engagement (p. 572). Unfortunately, due to financial situations or other responsibilities, many students must maintain employment while enrolled in their college courses, which can negatively affect engagement in their courses. In addition to personal issues, distractions, and navigating the college lifestyle, "students' lack of enjoyment of a course/field of study" can also negatively affect levels of course engagement (Benken et al., 2015, p. 20). Therefore, instructors and administrators at the college level should be aware of the challenges that could follow first-year students into the classroom.

At-Risk First Year Students in College

Some first-year college students are also considered at-risk. It is, first, important to understand who is considered at-risk. For the current study, students are considered at-risk if they have a greater chance of not completing their college courses to earn the degree they desire and ultimately "pose a greater attrition risk" at their college or university (Alward, 2012, p. 43). Students who are required to take developmental courses (or who need remediation) fall into this category. Placement into such courses is often directed by a student's performance on examinations like the ACT or SAT (Benken et al., 2015; Fong et al., 2015). Further, some groups

are generally more at risk than others. Women, when compared to men, and minority students, when compared to Caucasian students, are considered more at risk, especially in STEM related courses and fields because of their "lack of math background" (Betz & Hackett, 1993, p. 18). This is related to and perpetuated by the gender and racial stereotypes that are discussed in the section that follows.

It is also important to distinguish between the terms *at-risk* (synonymous with endangered, vulnerable, threatened) and *remedial* (synonymous with corrective, counteractive, curative) because they are not synonyms themselves. On one hand, students are considered at-risk if they are deemed to have a greater chance of not completing their courses to earn the degree they desire. On the other hand, the term remedial refers to courses that focus on reviewing or mastering pre-college content. Therefore, at-risk refers to a student's chance of non-completion and remedial refers to non-mastered pre-college content. At-risk students placed into a remedial course is what categorizes them as remedial students. Thus, at-risk is not the same as remedial, but a relationship between them exists.

At-risk college students are statistically "less likely to complete a degree; specifically, only 52% of remedial students attending four-year universities … and 28% of those attending two-year colleges" will complete their degree program (Benken et al., 2015, p. 15). A misalignment between high school graduation requirements and minimum college requirements likely contributes to the ill-preparedness of these students (Benken et al., 2015). Another reason for these low completion rates is that the students enrolled in developmental courses "lack not only the necessary content knowledge but also the appropriate learning and motivational strategies to succeed" (Fong et al, 2015, p. 56). Ultimately, these reasons lead to higher failure rates in these courses (Zientek et al., 2019).

Importance of Mathematics

The importance of mathematics as an academic domain will now be established. Mathematics is considered a gatekeeper subject and is one that college students are required to obtain a certain level of proficiency in to obtain a degree (Nguyen, 2015; Thiel et al., 2008: Alves et al, 2016). Courses like College Algebra, which is the focal course of the current study, are traditionally "required for several majors…A fear of math and lack of interest in algebra [can cause] many students to avoid the course as long as possible, often delaying their progress toward a degree" (Thiel et al., 2008, p. 46). This can often be difficult for students to achieve who require remediation, which adds additional credits/courses, time, and cost to their plates. Because of this, these students are less likely to finish their degree (Benken et al., 2015). Additionally, many factors are connected to issues of learning and teaching of mathematics. These include relevancy, utility, associated anxiety, self-efficacy, and the preconceived or experienced difficulties that people link to mathematics. Kapasa et al. (2015) focused on the issues of teaching and learning of mathematics, particularly approximation, and found that instructor and student attitudes are a major contributing factor. The factors that were previously listed can lead to other difficulties like lower levels of student achievement and motivation. In particular, self-efficacy, which is the focus of the current study, can affect student achievement and motivation in mathematics (Alves et al., 2016, Pajares & Graham, 1999; Zientek et al., 2019). Ultimately, a focus on mathematics supports instructors and the field of mathematics education to understand how to better support at-risk students as they complete their desired program.

Not only is mathematics an important academic domain and poses problems within academia, but it is also problematic as a discipline and poses non-academic issues for students,

which include stereotypes on race and gender. There are many well-known stereotypes linked to mathematics such as female students are not good at math (Alves et al, 2016) or that male students and Asian students are good at math (Jo, 2012). Stereotypes such as these can affect students inside and outside of the classroom and have particularly significant effects on female and minority students. First, "stereotype threat research indicates that minority students' cognitive strategies can be impaired when ethnic identity and a stereotype of being academically underprepared (e.g., minority students are bad at math) is made" (Fong et al, 2015, p. 60). A second example is when female students are successful with a mathematics task or exam, but the stereotype still causes them to feel inferior in their abilities to males (Alldredge & Brown, 2006; Betz & Hackett, 1983). Third, stereotypes that may be thought of as positive, can do more harm than good to those that are the subject. Prindle (2014) reported that:

> Adults often bestow the 'gifted' label on children who excel in math. This works to the detriment of the children who are labeled as gifted as well as those who are not. Children who believe that mathematical ability is fixed are vulnerable to losing confidence in their ability when confronted with difficulty (Joet et al., 2011). (p. 48)

In addition to racial and gender stereotypes, there is also the common belief that the ability to do mathematics is genetically determined. Hekimoglu and Kittrell (2010) report that some students use their predominantly negative experiences with mathematics to fuel the belief that "they were simply not suited for mathematical thought-born without a '*math gene*'" (p. 301). This belief essentially eludes individuals into thinking that people are born with or without the ability to be successful with math. Although this is a myth, it is still a belief that people carry with them in an effort to explain the difficulties they have encountered with the subject. This sense that mathematics ability is genetically determined coheres with and may explain why so

many people describe themselves as "not good at math" or "not a math person." Thus, the focus on mathematics is important because of the controversial issues it is associated with inside and outside of the classroom for all ages.

Developmental Mathematics Courses in College

Before discussing the nature of developmental mathematics, it is important to distinguish between the terms *remedial* and *developmental*. According to Higbee (1993), the term *remedial* has a medical, or curative, connotation. This contrasts with the way *developmental* is defined, which is a "focus on the process of learning as well as the content to be mastered" (Higbee, 1993, p. 99). In thinking about the feelings associated with these terms, *remedial* sounds negative and *developmental* sounds more hopeful and positive. The use of these terms, however, may not make any difference to the affected students. Both are used in the literature, but they should not be used interchangeably. For the duration of this document, the term *developmental* will be used.

Developmental mathematics courses, sometimes noncredit bearing, are designed to help students who do not place into college level mathematics, as deemed by the university and placement examinations. Many of these courses or programs work to "provide students with genuine mastery experiences and reduce excessive math fears. [They] may enhance students' math self-efficacy and, ultimately, expand their range of perceived career and academic options" (Lopez & Lent, 1992, p. 11). Further, these courses try to bridge the gap from high school to college level work and often try to level the playing field for students who may not have had the same opportunities as others (Fong et al., 2015).

Even though developmental level courses have been integrated into the selection of courses offered at many institutions, some question how successful these courses are in helping

11

students achieve their goals in college. The study by Noble and Sawyer (2013) considered just that. The authors stated:

> Most research to date has compared the overall success of developmental students as a group with that of non-developmental students; a typical finding is that developmental students are not as successful in the long term as non-developmental students. We ask a different question: "Does taking developmental courses benefit students at all, in the sense that they are more successful than they would have been if they had not taken developmental courses?" (Noble & Sawyer, 2013, p. 4)

This, however, is a difficult question to answer because the students can only take one of these courses, whether it is developmental or not. So, it is important to note that Noble and Sawyer (2013) made comparisons between students in developmental and non-developmental courses but who had "the same ACT Test scores, enrollment status, and who enrolled in similar institutions" (p. 57). The authors found, from consideration of "ACT Test student records and college outcomes data for 118,776 students who first enrolled in one of 75 postsecondary institutions," that certain students enrolled in developmental courses did experience some benefits, even though the group as a whole did not seem to do as well as those enrolled in non-developmental courses (Noble & Sawyer, 2013, p. 8). Students who seemed to benefit more from developmental coursework included those who were better prepared academically prior to college according to ACT scores, those who earned higher grades in the developmental course they enrolled in, and those who had part-time enrollment status (less than 24 credit hours in one academic year). Therefore, there is some support for the claim that developmental courses can be supportive of some students' success in college.

Developmental courses look different than a standard college course. While many standard courses assume mastery of all prerequisite concepts, developmental courses work to fulfill these prerequisites by "expand[ing] the academic skills of students" (Hall & Ponton, 2005, p. 26). These skills include critical thinking, communication in various forms, and problem solving, among others, for the subject(s) in which the student is enrolled (Higbee, 1993). Ultimately, students who enroll in developmental courses are not believed to be ready for traditional college coursework (Zientek et al., 2019). One reason for this is that the curriculum of these courses is comparable to what students did not master, but have seen, in high school. Although, Higbee (1993) advises that it should not be a complete reiteration of the material because the students do not want that, given the position they are in. A second reason comes from Benken et al. (2015) who points toward misalignment of high school and college curricula because, in the case of mathematics, "students typically need to complete Algebra I, or the equivalent" to graduate high school, which does not align with how "colleges expect students to complete at least two courses beyond that and maintain a fundamental understanding" (p. 15). So, some students were not actually exposed to the material that they were determined to not have mastered. Regardless of the reason for enrollment, development of skill and content are the foci of these developmental courses, as well as "the development of the 'whole learner'" (Higbee, 1993, p. 100). These courses should also require:

> A high degree of structure…as a source of support for developmental education students. The course syllabus must outline goals and objectives, list assignments and test dates, and clearly state attendance policies and the grading scheme. Faculty members must adhere to deadlines, start and end class on time, return assignments promptly and maintain scheduled office hours. (Higbee, 1993, p. 100)

Developmental mathematics courses should follow the general guidelines outlined by Higbee (1993) above to better support enrolled students; the treatment course described in the current study does. Overall, instruction must be sensitive to students and their needs, and "teachers of underprepared students must ignite excitement" (Higbee, 1993, p. 103).

Where developmental courses are more frequent at the community college level, they also exist at four-year institutions, like the one that is the setting of the current study. This is true because "public two-year, postsecondary institutions of education …provide vocational degrees, transfer programmes to four-year universities, and access to education for all students," and so more students that need these courses begin their studies here (Zientek et al., 2019, p. 183). No matter what type of institution a student is enrolled in, if they do not succeed in their developmental mathematics course, then "they are very likely to withdraw from college within 1 year" (Cafarella, 2016, p. 13). Even though these courses are designed to support these students, developmental course requirements may actually be hindering student enrollment (Cafarella, 2016) because of additional coursework, time, and cost (Benken et al., 2015). In order to help students succeed, accelerating and compressing these courses has become a conversation amongst those who work in higher education because some institutions have experienced success with these modifications (Cafarella, 2016).

Social Cognitive Theory

Before delving into a discussion about self-efficacy itself, it is important to discuss the theory it stems from. Initially called Social Learning Theory, Bandura expanded his work and coined Social Cognitive Theory to explain individuals' behaviors in learning and further highlight the role of cognitive influences. This theory can be applied to many fields including education. According to Bandura:

14

Human agency operates within an interdependent causal structure involving triadic

reciprocal causation (Bandura, 1986). In this transactional view of self and society,

internal personal factors in the form of cognitive, affective, and biological events;

behavior; and environmental events all operate as interacting determinants that influence

one another bidirectionally. Reciprocity does not mean that the three sets of interacting

determinants are of equal strength. Their relative influence will vary for different

activities and under different circumstances. (1997, p. 6)

These interactions between person, behavior, and environment work together to explain

behaviors in learning. An influential construct of Social Cognitive Theory is self-efficacy, which

plays a major role, and will be discussed in the following section. Its role is one that can affect

the degree to which or whether at all an individual participates in specific activities or engages in

certain domains. Self-efficacy fits centrally within the broader theory because the construct

develops from how an individual makes sense of his/her environment (which includes internal

and external influences) in order to determine beliefs about his/her own abilities, ultimately

affecting behavior. Self-efficacy beliefs, which are specific and not general, are constructed

based on an individual's observations of oneself and observations and/or interactions with others

within an environment or under certain conditions. Individuals use these specific constructed

beliefs about their own abilities to determine behavior, or whether they will engage with the task.

Thus, "people guide their lives by their beliefs of personal efficacy" (Bandura, 1997, p. 3).

Self-Efficacy and Self-Efficacy for Mathematics of At-Risk Students in College

Self-efficacy influences many things in mathematics, including achievement and

motivation (Alves et al., 2016, Pajares and Graham, 1999; Zientek et al., 2019). It has also been

shown to influence and predict academic performance (Katz, 2015; Pajares & Miller, 1994) and

more specifically GPA (Fong et al., 2015), college major choice (Usher & Pajares, 2008), and career choice (Bong & Skaalvik, 2003; Usher & Pajares, 2008). The importance of the construct is held by researchers from many fields, both related and not related to mathematics. Usher (2016) noted that most of the research on self-beliefs has been done with mathematics but that it is important to explore the effects in other domains more than what has currently been done including "social studies, civics, music, art, engineering, history, science, and language learning" (p. 156).

Self-efficacy, according to Bandura, is defined as the "beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments" (1997, p. 3). This definition is cited and used by many researchers in the field today. There are four main sources of self-efficacy: (a) mastery or performance accomplishments, i.e. having past success with a task or domain promotes self-efficacy; (b) vicarious experiences, i.e. seeing others who are similar to oneself being successful in a task or domain promotes self-efficacy; (c) verbal persuasion, i.e. positive feedback from others in relation to the task or domain promotes self-efficacy; and (d) emotional arousal, i.e. positive feelings associated with a task or domain promotes self-efficacy (Bandura, 1977). Mastery experiences has been the strongest predictor of self-efficacy (Bandura, 1977). In recent research, Zientek et al (2019) studied the self-efficacy of community college students who were required to take developmental mathematics courses. The results indicated that Bandura's (1977) four sources were influential, with mastery experiences being most influential. Additionally, Smith (2010) stated that students in developmental courses "were typically uncertain about their goals and had low self-efficacy on some academic tasks" (p. 49).

People have self-efficacy for many specific tasks and academic domains, but not generally. When researchers work to understand and measure self-efficacy, Pajares and Miller (1994) cautioned:

> Self-efficacy must be specifically rather than globally assessed, must correspond directly to the criterial performance task, and must be measured as closely as possible in time to that task. These guidelines are seldom followed, and so the mismeasurement of self-efficacy is a recurring theme in educational research, often producing poorly defined constructs, confounded relationships, ambiguous findings, and uninterpretable results. (p. 194)

Since self-efficacy is subject-specific, it does not only pertain to mathematics.

Let us briefly consider how it affects other subjects because students do not just take mathematics courses, they take courses for other subjects for which they also have varying levels of self-efficacy. For example, a student may have low self-efficacy toward mathematics, but high self-efficacy toward foreign language. Self-efficacy in other subjects affects students in comparable ways as it does for mathematics, and "is an important motivational variable in predicting a variety of achievement-related behaviors" (Linnenbrink-Garcia et al., 2018, p. 191). Linnenbrink-Garcia and colleagues focused on pharmacology within the science domain, specifically considered GPA as an achievement-related behavior, but also considered self-efficacy's relationship to persistence behaviors like course completion and career intention, as well as task value. Upon conclusion of data analysis, the researchers had evidence that their intervention (a two-week summer program) supported higher levels of self-efficacy after eight months, but their data was not consistent with self-efficacy's predictive power, which they noted as surprising in some cases. They, however, still supported self-efficacy's predictive nature. It

has also been shown that "motivation and self-efficacy of individuals are directly proportional to their success in biology" (Ates & Saylan, 2015, p. 91). The same can be said for other subjects, including engineering (Ernst et al., 2016). In addition, Larson et al. (2015) considered self-efficacy's predictive power on retention and found that "mathematics/science self-efficacy from the first semester at a university significantly contributed to graduation status four to eight years later" (p. 15). Graduation status in this study refers to whether participants graduated or not. Self-efficacy has been shown to predict many educational outcomes in relevant literature including academic achievement (Fong et al., 2015; Katz, 2015; Usher & Pajares, 2008; Yurt, 2014; Zientek et al., 2019) and more specifically GPA (Fong et al., 2015), college major choice (Usher & Pajares, 2008), and career choice (Bong & Skaalvik, 2003; Usher & Pajares, 2008). Thus, the importance of studying self-efficacy is something that is recognized by researchers in domains both related and not related to mathematics.

The relationship between self-efficacy, motivation, and math anxiety. Connected to self-efficacy are issues of math anxiety and motivation. A generally accepted relationship in the field between these is that self-efficacy and motivation are positively correlated, while self-efficacy and math anxiety is negatively correlated (Ates & Saylan, 2015; Watts, 2011). Math anxiety is "a state commonly described as approaching mathematics with trepidation due to related feelings of weakness, dependency, and frustration" (Hekimoglu & Kittrell, 2010, p. 301). It is important to consider math anxiety for at-risk students because many of them, in my experience, express they have anxiety toward mathematics and report negative experiences with the subject in the past. These feelings can negatively affect their motivation and self-efficacy toward the subject. A factor also linked to math anxiety, especially toward exams, is the classroom environment which is influenced and created by the instructor (Ates & Saylan, 2015).

Self-efficacy beliefs and age. We begin with research on pre-college students. Many researchers have found that self-efficacy in younger students, generally those in elementary school, is still developing and so their responses to questions about their self-efficacy may not be as accurate if they answered the same questions when older (Joet et al., 2011). Further, younger students may, inadvertently, exaggerate their self-efficacy beliefs so that they appear higher than they actually are (Panaoura et al., 2009). Students in elementary grades take a single or few experiences of success or congratulatory words from a teacher, parent, or another student and use those to fuel their self-efficacy beliefs.

In secondary grades, researchers have found that these students tend to have lower and more realistic self-efficacy beliefs because of more experiences and awareness. Further, Panaoura et al. (2009) stated:

> It seems that students' sense of efficacy diminishes somehow when they compare their abilities with classmates, and even more in relation to their mathematical performance as it is revealed by their final grades at mathematics.… This decrease can be explained by the difficulties that are increased in secondary education, … On the other hand, it can be explained by the construction of a more precise self-image when students become older. (p. 723)

Therefore, as students age, they develop more stable self-efficacy beliefs for mathematics. This can, however, have a negative effect on their motivation and math anxiety.

College students are more like secondary students in terms of their self-efficacy beliefs. Self-efficacy has been found to be quite influential and "more predictive of mathematics performance in college students than other factors, including the mathematics self-concept, perceived usefulness of mathematics, prior experience with mathematics, or gender" (Kung,

2009, p. 389). Generally, students with a higher level of self-efficacy seem to persevere when faced with challenge and perform higher than those with lower levels of self-efficacy in mathematics (Gumml, 2013). And, for those enrolled in developmental courses, their self-efficacy is generally lower (Smith, 2010) which impacts how they approach, or often avoid, the subject of mathematics.

Self-efficacy interventions. For the purposes of this study, an intervention is defined as a purposeful act(s) by a researcher in order to enact some sort of desired change when measured against a control. In the current study, the acts are situated within the interactions of students, instructor, and mathematics, and the desired change is improvement in students' self-efficacy toward mathematics. Interventions to help students build a positive self-efficacy in mathematics have targeted ages from elementary to college level and have been diverse in character.

Pre-college intervention work that was considered in this review has focused on mid-upper elementary through high school. This focus on higher grades (and not early elementary) is likely because younger students' self-efficacy beliefs can be exaggerated and unrealistic (Joet et al., 2011; Panaoura et al., 2009). About a fairly equal number of studies focused on college intervention work on self-efficacy in mathematics. Some of these studies, no matter the participants' age, have focused on the construct's effects on student achievement, some focus on teachers' orientation toward students, some focus on teaching study skills strategies, and some focus on changes in curriculum or course structure.

In terms of the length of time that these studies were framed, pre-college work has varied from four weeks up to a full academic year, but more commonly lasting three months or less. For those that focused on the college-age population, intervention work varied from one hour to one full college semester, with more work being done in time frames of one month or longer. It is

important to note that much of the work reviewed for this study did not appear to assess longitudinal effects of the intervention. Longitudinal studies, for this review, are those that intentionally follow-up with participants at least once after the conclusion of the intervention, whether it be weeks, months, or years later to assess the existence of long-term effects. The closest that studies in this review came to this was assessing the effects of a one-hour intervention six weeks later or following up at the end of the semester after a four-week intervention. Further, all but one study (whose focus was statistics) reviewed for this literature review focused on self-efficacy in mathematics.

If we consider subjects related to mathematics, in particular those that are considered STEM (Science, Technology, Engineering, and Mathematics), more work about longitudinal effects can be discussed. In particular, the work by Linnenbrink-Garcia et al. (2018) considered longitudinal effects of an intervention, in the form of a two-week summer program (between the first and second years) implemented to support STEM persistence in first-year undergraduate students who were enrolled in an early chemistry course. This summer program focused on "fundamental concepts in pharmacology (week 1) and drug treatments for four specific diseases (week 2)….The majority of the course involved inquiry-based and active-learning activities, with less than one hour/day devoted to lecture" (Linnenbrink-Garcia et al., 2018, p. 185). Upon conclusion of the intervention and a follow-up survey, the authors concluded that the "summer program was associated with higher self-efficacy eight months later" (Linnenbrink-Garcia et al., 2018, p. 191). Thus, this study illustrates an example of longitudinal work that can be done to measure self-efficacy in STEM upon completion of an intervention.

Another study by Raelin et al. (2014), which did not utilize an intervention on self-efficacy, focused "on the retention of women in the STEM fields by inquiring whether and to

what extent their retention, as well as that of their male counterparts, is due to changes in their self-efficacy" (p. 3). Undergraduate participants in engineering programs from four universities completed surveys during their second, third, and fourth years, with a collection of data related to graduation at the conclusion of the fifth year. Three scales on self-efficacy were considered. First, work self-efficacy, or "students' beliefs in their command of the social requirements necessary for success in the workplace," was measured with a scale designed by one of the researchers (Raelin et al., 2014, p. 12). Second, career self-efficacy was "students' confidence in accomplishing a range of tasks having to do with their prospective career" and was measured with a second scale (Raelin et al, 2014, p. 17). Third, academic self-efficacy was "one's confidence in accomplishing a set of tasks having to do with the engineering major" and was measured with a third scale (Raelin et al., 2014, p. 18). Upon conclusion of the study, Raelin et al. (2014) reported:

> Academic self-efficacy and contextual support in all time periods were found to be critical to retention. Contextual support was found to be particularly important to women and appears to serve as an inducement to stay in school and in the major. Work self-efficacy, developed by students between their second and fourth years in school, is also an important factor in retention, though it is strongly tied to the participation by students in co-op programs. Career self-efficacy did not play a unique role in the change analysis. (p. 38)

Therefore, this study illustrates the importance and influence that long-term effects of self-efficacy have in STEM.

All interventions on self-efficacy in the two sections that follow focus on increasing student self-efficacy toward mathematics but are divided into four groupings based on their

22

method of influence: (a) a focus on student study and motivation skills, (b) a focus on direct teacher influence, (c) a focus on changes in course structure, and (d) a focus on a change in curriculum. I begin with research on interventions for pre-college students and then shift to studies of college students. Each intervention will be described and then a synthesis about the interventions' perceived or measured level of success, limitations, and gaps in the literature will follow.

Interventions for pre-college students. For studies that focused on pre-college students, the various interventions focused on three out of the four listed methods of influence: student study and motivation skills, teacher influence, or course structure. We start with those that focused on student study and motivation skills. First, the 8-week intervention detailed by Dennis (2013) involved three treatment groups for students in the third and fifth grades. Each of the intervention groups received one of the following interventions: (a) cross-age peer tutoring in mathematics in which a fifth grader was paired with a third grader to tutor in multiplication, (b) self-efficacy skill building in the form of 30 minute lessons in which four concepts (time-management, goal setting, studying, and test taking) were taught over the period of 8 weeks, and (c) a combination of these two treatments. Dennis (2013) found "that third-grade students, who participated in the combined treatment group…had higher levels of self-efficacy at the final measurement…At the fifth-grade level…the peer-tutoring group had the highest levels of self-efficacy at the final measurement." (p. 80).

Second, Katz (2015) worked with sixth grade students for an academic year on strengthening their abilities in "goal setting including reinforcement; skill and strategy acquisition, and open reflection tasks on their self-efficacy to learn mathematics" (Katz, 2015, p. 46). Goal setting began daily, then weekly, then monthly, then every three months. Students

made new goals after one was achieved and they met with teachers to discuss goals. Students were also taught various math concepts and skills, provided with experiences of mastery, and then asked to complete 20 unstructured free-write reflections. This study focused on providing a student profile rather than a self-efficacy score. At the end of the intervention, Katz (2015) found improvement in student self-efficacy, and improvement among other foci. The authors credited the student profiles in helping to diagnose self-efficacy prior to the intervention.

Third, a study by Prindle (2014) taught "lessons on the specific contributors to self-efficacy: exercising optimism, enhancing emotional literacy, and engaging intrinsic motivation" to fourth grade students via Social Emotional Learning (SEL) techniques (p. 4). This was done to improve students' self-efficacy and performance in mathematics over a period of three months. Data was collected in the form of observations, pre- and post- achievement test, surveys, daily student journals, and parent interviews. Noticeable differences in self-efficacy and performance in mathematics was achieved after the intervention. Though these three studies focused on study and motivation skills, they also targeted the learning of content as well.

Direct influence from teachers can also be used to influence student self-efficacy toward mathematics. For example, Siegle and McCoach (2007) studied whether staff development on self-efficacy strategies to use during instruction could affect fifth grade students' self-efficacy for mathematics. This training lasted two hours for teachers who participated in the intervention group. In particular:

> The training specifically focused on teacher feedback, which included teachers complimenting students on their abilities and the skills they acquired; goal setting, which included activities designed to draw students' attention toward their successful

performances; and modeling, which involved students observing fellow students successfully completing similar tasks. (Siegle & McCoach, 2007, p. 282)

Teachers and their students from the intervention and control groups participated in a four-week mathematics unit following the training session. The results indicated an increase in student self-efficacy for mathematics for the intervention group.

Researchers have also used changes in course structure as an intervention. In one study, Perry and Steck (2015) used iPads in high school geometry class for one semester. One class used the iPads, while a second class did not. The goals were to increase student engagement, improve scores, and increase self-efficacy and self-regulation by merging this technology "into course instruction as an add-on support tool" for learning in the classroom (Perry & Steck, 2015, p. 130). The use of iPads is considered a change in course structure for this review because the added resource to support students in learning geometry concepts was a component incorporated into the original curriculum and instruction. It served as an added resource because of its capabilities to "[engage] students in completing problems using iPad applications" (Perry & Steck, 2015, p. 130). The results of its use were scrutinized. After the intervention, "students in the iPad-using classroom experienced lower levels of geometry proficiency scores, higher levels of off-task behaviors, and similar levels of self-efficacy and metacognitive self-regulation compared to the non-iPad group" (Perry & Steck, 2015, p. 122). Thus, the intervention in this study was not successful, even though there were no significant decreases in self-efficacy toward mathematics.

Interventions for college students. For studies that focused on college students, the various interventions also focused on three out of the four listed methods of influence: student study and motivation skills, course structure, or curriculum changes. We begin with those that

focused on student study and motivation skills. Cordero et al. (2010) implemented two interventions focusing on study skills to increase college students' self-efficacy toward mathematics. One group, which the authors referred to as the "control group," received a performance accomplishment and the other, "experimental group," received the performance accomplishment treatment paired with belief-perseverance techniques. The performance accomplishment treatment, which refers to the *mastery experiences* source of self-efficacy, was a chosen self-efficacy enhancing number series task in which "participants were given 12 min to solve 18 problems by determining the pattern of each series and identifying the next two numbers in the sequence" (Cordero et al., 2010, p. 367). Additionally, the authors defined belief-perseverance as "the tendency for people to continue to accept their thoughts and opinions as true in the face of disproving facts" (Cordero et al., 2010, p. 364). The authors were interested in whether, after the intervention, students would maintain the beliefs they wrote about in the experimental writing prompt that is described next. All students completed the same math task and one of two writing prompts. The experimental prompt was about student deservingness of a math/science scholarship and the control prompt was about student deservingness of a travel scholarship. Following the math task and writing prompt, a posttest scale was administered. All of this occurred in a one-hour session and then a follow-up survey was emailed six weeks later to measure lasting effects. The experimental intervention produced more positive gains than the control because the students retained their belief in their argument for why they deserved the scholarship and could succeed in mathematics and science courses. This study suggests that "if individuals could construct, for themselves, a rationale about why they can succeed in math and science, they might believe the rationale and experience a corresponding increase in their math

self-efficacy" (Cordero et al., 2010, p. 364), which was demonstrated in the results and discussion of the study.

Perhaps because college instructors have more freedom in structuring their courses than K-12 teachers, more studies, like the four described next, have studied the effects on self-efficacy of changes in course structure in college classrooms. First, Brewer (2009) studied two groups of college students in multiple sections of a one-semester college algebra course. The control group completed homework via textbook; the treatment group completed homework via an online program to see if changes in self-efficacy and/or mathematics achievement could occur. Upon completion, Brewer reported increases in student self-efficacy for mathematics in both groups, but no significant differences between them or in their levels of achievement. The author determined "that using online homework as part of a college algebra class to facilitate learning and mathematics self-efficacy is at least as effective as using textbook homework" (Brewer, 2009, p. 127). Also, Brewer (2009) discussed, from report of other works, that perhaps factors like the level of advancement of the online program, the length of the treatment period, and/or the amount of time students spent on the online program may have affected the results. However, it was determined that the online homework seemed to be more beneficial for underprepared students (Brewer, 2009).

Second and third, Hodges (2008) and Hodges and Kim (2010) also worked to influence student self-efficacy for mathematics using course structure. The intervention in both studies was the use of email in a one semester online college algebra and trigonometry course; messages to students were sent for four weeks and 11 weeks, respectively, to provide encouragement and positive affirmation of progress made. For example, a personalized message that a participant might receive after submitting a quiz would say, "Very nice! You correctly answered all but one

question on Math 101 quiz [quiz number]. Set aside enough time to study math next week and try for a perfect score" (Hodges, 2008, p. 270). These messages were designed to address the verbal persuasion source of self-efficacy. The treatment group received messages like this, while the control group received messages that were more neutral in nature. Additionally, in Hodges and Kim (2010), a second treatment group received these messages, but without a personalized greeting. An example of a message that a control group participant might receive would say, "Coin return lockers are available at the Math Emporium to keep your personal items safe. The lockers accept quarters only and are located in the front lounge" (Hodges, 2008, p. 285). Control participants in Hodges and Kim (2010) would only receive messages that directed them to a survey four different times. The studies similarly reported that, "contrary to expectations, the email messages did not lead to any positive change either in self-regulation or in self-efficacy" (Hodges & Kim, 2010, p. 216). It is important to note that Hodges (2008) reported initial increases in self-efficacy followed by decline. In addition, both studies observed a relationship between self-efficacy and achievement.

A fourth intervention, which was implemented by Valencia-Vallejo et al. (2018) studied the effects that motivational scaffolding had on self-efficacy in an online learning environment for two months. The experimental group received the scaffolding and the control group did not. The interactive scaffolding, which had four stages: Assessment, Goal Formulation and Planning, Learning Environment Execution and Monitoring, and Self-Evaluation and Final Reflection, was implemented into the online program. In particular, during the first stage (Assessment) students were asked to reflect on their knowledge and efficacy in mathematics and took a test to determine actual knowledge. During the second stage (Goal Formulation and Planning) students set a goal, determined the likelihood of achieving that goal, and made a plan to achieve that goal.

In the third stage of the scaffolding (Learning Environment Execution and Monitoring) students monitored their progress and revised, if necessary, their plan to achieve the goal they previously set. During the fourth stage (Self-Evaluation and Final Reflection) students evaluated and reflected on their efficacy and how well they worked toward their goal. Upon conclusion of the study, the scaffolding treatment had positive effects on students' self-efficacy and achievement.

Other researchers have explored changes in curriculum to influence college students' self-efficacy toward mathematics. Olani et al. (2011) developed a one semester reform-based introductory statistics course to measure self-efficacy and other beliefs of college students. Elements of the course included those that reflected the GAISE College Report including "goals, content, use of technology, and assessment" (Olani et al., 2011, p. 52). This reform-based course used active learning techniques, a combination of large lectures and small-group sessions, students received lecture slides online to review before class, and students had the ability to engage in online discussion boards about the course. Upon conclusion of the course, observed increases in student self-efficacy were reported, especially for those who recognized support from their instructor. Another example included an action-research intervention by Hekimoglu and Kittrell (2010). Students from college algebra, pre-calculus, and calculus courses watched a documentary on mathematics and wrote a reflection prior to a whole class discussion, which focused on a specific mathematician and the art of proofs (Hekimoglu & Kittrell, 2010). The authors described the effects of the intervention as possibly decreasing math anxiety and possibly increasing math self-efficacy. Evidence of this came from student written reflections, end of course evaluations, in-class discussions, and observations. The authors also report "a significant rise in students' grades and a reduction in student drop rates" since the implementation of the intervention (Hekimoglu & Kittrell, 2010, p. 324).

Considering the results of the work done by each of these researchers in terms of their chosen method of influence, no matter the age of the treatment population, those that focused on students' study and motivation skills were all successful, the one study that focused on teacher influence was successful, those that focused on course structure were successful in about half of the studies, and those that focused on curriculum changes appeared to be successful (although one vaguely concluded possibly successful while the other did not have a control group). Further, a greater number of studies focused on implementing an intervention that involved a change in course structure than any other type. It is interesting that the studies grouped together with this focus seemed to have the least success of the four groupings, even though it seemed to be the most popular form of intervention. The strongest, positive results appeared to come from those that focused on work with students' study and motivation skills, which was the second most popular form of intervention. Upon discussion of their results, many of the researchers also indicated limitations of their work.

Limitations of past intervention work. The studies discussed in this review have indicated various limitations of their work. The most commonly mentioned limitations were small sample size which affected some of the studies' generalizability, short time frame of some studies, participation issues (i.e. participants withdrawing or issues with Hawthorne effect), missing data for various reasons, and demographic-related issues. Other limitations included the inappropriateness or ineffectiveness of the chosen intervention, difficulties associated with participant observation, choice of technology used in one study, no control group used in one study, difficulties experienced from using multiple instructors (i.e. inconsistent pedagogy practices), misalignment between measures and intervention, and the confliction identified from one study of a teacher playing a researcher role. These limitations were openly discussed by the

30

researchers whose work was included in this literature review. It seems that the most common limitations for work in this field are related to small samples of participants and quite short-duration study time frames. Therefore, future work in this area should focus on having larger group(s) of participants to be able to generalize results and should cover a greater length of time to understand longitudinal effects.

This review of interventions at the K-12 and college level illustrates the variety of approaches researchers have used to positively influence students' self-efficacy toward mathematics and the limitations these researchers have encountered while attempting to do so. The current study is another example of an intervention or treatment with the goal of positively influencing student self-efficacy, yet with a different design. If the current study were classified using the four approaches listed above, it could be classified under all four. Further, important gaps in the field were identified in the literature, which the current study seeks to fill. The first is that "few intervention studies investigate how to improve math self-efficacy effectively" (Cordero et al, 2010, p. 363). For the articles reviewed, many more of them focused on trying to understand self-efficacy, its relationship with other constructs or factors, and its effects on students, while fewer focused on improving academic conditions to positively affect or influence growth in self-efficacy. The second was identified by Zientek et al. (2019) which is a "gap in understanding self-beliefs of developmental mathematics students" in particular (p. 186). For the articles read for this review, few of them specifically focused on the at-risk student population enrolled in developmental mathematics courses, as opposed to more for college students in general.

CHAPTER 3: METHODOLOGY

The purpose of the current study was to measure and understand changes in student self-efficacy for mathematics over the period of one semester while participants were enrolled in a special treatment section of a developmental algebra course that I taught. The need for a study such as this is supported by the literature; the field is missing more studies that seek to improve students' self-efficacy toward mathematics and/or those that work to understand such beliefs of college students in developmental mathematics courses. It is important to improve students' self-efficacy toward mathematics because mathematics is a domain that most people use daily in many tasks. It is important to understand the self-beliefs of college students in developmental mathematics courses because these are often the students that have lower self-efficacy, which can be indicated by their past experiences involving fewer instances of success. An understanding of student self-efficacy could lead to better ways to improve levels of the construct. This study examined the effects of trying to positively influence the self-efficacy of at-risk first year college students in a developmental algebra course. The perceived effects of the course as a treatment, which is described in sections to follow, on students' self-efficacy toward mathematics were measured with a standard instrument at the start, middle, and end of the semester. Those results were compared to two sets of student interviews, periodic student journals, student grades on course assessments, and an instructor perspective in order to determine if/how students' self-efficacy changed possibly from their enrollment in the treatment course and how students made sense of that change, if it occurred. As explained later, I planned for specific steps to separate myself from the data collection.

In my experience, students come into the developmental algebra course with feelings associated with negative self-efficacy for mathematics. The materials that I integrated into the

course were designed and/or chosen from various sources to encourage students to see connections between mathematics and its value in the world, help students master the objectives that are required by the university, and help them see and find success with mathematics. It was my hope that these sparked positive changes in their self-efficacy toward mathematics as well. It is supported and advised by the literature that providing students with experiences of mastery or success can increase their self-efficacy (Betz & Hackett, 1993; Zientek et al., 2019). In the past, these materials have been positively received by students who were considered at-risk and who traditionally appear to have negative self-efficacy toward mathematics.

Setting

The setting of the current study was at a mid-size public university in the rural Midwest. The university serves just under 15,000 students each year. Roughly 190 degree programs are offered at this institution including two-year programs, four-year programs, and graduate programs.

Available at this university is a program for at-risk students who qualify for developmental placement in reading, English, and mathematics, called the C4 Scholar Program. Only one cohort and one section of each of the three courses are run each year. If qualified students choose to participate in this program, they save time by being placed into courses that are one level above their placement (based on SAT scores), they save money from not having to take non-credit bearing courses, and they are given extra supports that they normally would not receive in a traditional mathematics course. These supports include instructor-led workshops, a study session built into student schedules, a learning community of peers, opportunity to revise work in the three core courses, and a team of instructors who communicate across disciplines frequently throughout the year in order to holistically support student progress. These supports

help to provide students with opportunities for success and help to persist through struggle. Further, the focus of this program is retention of the at-risk student population at this university, by instilling accountability and professionalism in the students.

Evidence and support for these factors was found by one of the C4 co-researchers in the form of a framework called the Oz Principle. After the pilot year (2016-2017), the Oz Principle (Connors et al., 2004) was incorporated into the C4 Scholar Program which focuses on accountability and professionalism through Above the Line and Below the Line behaviors. This transition if reflected in the article by Williams, Conley, Pavletic, and Weller (2020). Above the Line behaviors are ones that promote personal and professional growth and productivity and are outlined as four steps by the authors: See It, Own It, Solve It, and Do It. In the C4 Scholar Program, students are encouraged to see a problem or a challenge they have, take ownership of the problem or challenge, work towards solving it, and then do something about it. Some students may not make it all the way to the Do It step, but any gains are seen as positive in the program. Alternatively, Below the Line behaviors are ones that are opposite of those just described and examples include denial, finger pointing, waiting to see what will happen, wanting someone to tell them what to do or how to do it, etc. (Connors et al., 2004). Past years (2017-2018 and 2018-2019) have been dedicated to identifying and responding to Above and Below the Line behaviors. The beginnings of and evolution of the C4 Scholar Program and the development of materials are discussed in Appendix A.

Students in the cohort for the 2019-2020 academic year are one group of participants in this study. Data was collected from them for the duration of one fifteen-week semester, with a sixteenth week for final exams. These students were enrolled in a treatment section of a College Algebra course (coded as MATH 115) that covered content from MATH 110: Fundamentals of

Algebra (or Beginning Algebra) and MATH 115: Intermediate Algebra in one semester instead of the traditional two-semester sequence. The course met four days per week, Monday through Thursday, for 50-minute sessions, with two 75-minute workshop periods. The Monday workshop was a general study session for students to complete coursework. In the past, all three instructors (English, reading, and mathematics) took turns overseeing this study time. Recently, as was the case for Fall 2019, a past C4 Scholar Student was paid to oversee the study session each week. The Wednesday workshop was always covered by the mathematics instructor, who is the researcher for this study, and it was used for additional assistance and practice in the mathematics course only. Both workshop periods were on students' course schedules to ensure their availability and most students have taken advantage of these study periods in the past, especially when they were optional. Attendance was recorded at each class period and workshop. The description of the elements of the course, which the researcher considers the treatment, is provided after the section about participants. The classroom in which the course was taught was equipped with whiteboards, a computer with projector access, and a document camera, all of which were used during the current study.

The 2019-2020 term was the instructor's fifth academic year at the university and fourth academic year teaching a C4 cohort. While at the university, the instructor has taught 34 sections of various mathematics courses to date including many sections of beginning and intermediate algebra courses, two sections of a statistics course for teachers, one section of a geometry content course for preservice elementary teachers, three sections of a contemporary mathematics course which covers a broad spectrum of algebra and some statistics content for primarily social work students, one section of a math analysis for business course, one section of a trigonometry

35

course, and six sections of a quantitative reasoning course for professionals (i.e. nursing students, business students, and social work students).

Participants

Treatment group. Twenty-four first-year students participate as a cohort in the C4 Scholar Program each year. In the past, there have been approximately 60-80 students that qualified for placement into C4 each year, as was the case for 2019-2020 (although for 2018-19, numbers of students who did not participate were about equal to those who did participate). Qualifications for placement in the past focused on ACT scores (less than 14 for English, 15 for Math, and 17 for Reading), but recently have focused on SAT scores (less than 430 for English and Reading, 500 for Math). High school GPA was noted but was not a determining factor in course placement. All students who qualified were sent letters that explained the C4 Scholar Program. These students, then, determined whether they wanted to participate. The first 24 who responded were accepted. I confirmed with the program coordinator that no students were turned away from the C4 Scholar Program for the 2019-2020 academic year. Only 24 students showed interest, and all 24 were admitted to the program.

Recruitment. These students were recruited in person for the current study on the first day of class of Fall 2019. All students were 18 years of age or older. I introduced the current study, explained what their responsibilities as a participant would be, what they would receive in return ($40 in gift cards to the university bookstore – $20 given at the middle of the semester and $20 given at the end of the semester), that everything including their identity, decision to participate, and data would be confidential to me until semester grades have been submitted, and introduced the two student data collectors, one of whom was present. It was my hope that every student in the 2019-2020 cohort would agree to participate. In order to avoid undue influence

36

from me, the instructor, I left the room as the students decided whether they wanted to participate, and the student data collector distributed consent forms (see Appendix B) for the students to review. If a student decided to participate, s/he would sign the consent form, give it to the student data collector, and would then complete their first survey (*MSES* 1) and Journal 1. Students were also told that they could think about this decision for a longer period of time and take the consent form home with them if they wished. At the conclusion of the first week of class, there were 12 participants in the treatment group.

In hopes to have more participants, I extended the recruitment phase of the current study for a few reasons. First, I thought the students were overwhelmed with the big picture of the current study. Second, I thought the students were overwhelmed with their new environment, meeting new people, and having a new schedule of classes. Third, a major holiday weekend concluded the first week of class and I felt, from past experience as an instructor at this university, that many students did not settle into the semester until that holiday weekend was behind them. In fact, one of the C4 students decided to officially withdraw from the university over the holiday weekend because of homesickness (who was not one of the original 12 participants) and the participant pool fell to 23 students. And, so, at the beginning of the second week, I provided a simplified overview of my study again and explained my reasoning for recruiting them once more. I told them that their decision to participate should be made by the end of the week and that they would not hear about recruitment any further. The students were made aware, if they wished to participate, to turn in their consent form to the student data collector. At the end of the second week the final count of the treatment group stood at 19 participants.

To help the student data collectors keep track of submissions from individual participants throughout the semester, I created a blank data tracker sheet (see Appendix C). The student data collectors agreed to follow up with any students that were absent on a day that data was collected. This sheet was placed into a sealed envelope that only the student data collectors could see. It was never opened by me until final grades were submitted.

Based on previous experience with the C4 Scholar Program, I anticipated that it was unlikely that all 23 students in the course would complete the C4 Scholar Program. In the past, students have withdrawn from the university for various reasons including personal and academic. These students may not even complete the Fall semester of their freshman year. Upon conclusion of the Fall 2019 semester, other than the one student who left after the first week, no other students left the university or the course during Fall 2019. If this situation were to have occurred with one of the 19 participants, a plan was outlined in the Data Sources and Collection section of this document.

It is also important to indicate that ongoing social drama transpired among the treatment group throughout the second half of the semester, starting in Week 7. The details of this drama are provided in Appendix D and come directly from daily instructor journaling. Occasional disputes have occurred in past C4 Scholar cohorts, but none to the degree that was observed with the 2019-2020 cohort.

Control Group. Students who were not enrolled in the C4 Scholar Program, but who had the same incoming mathematics course placement ($n = 395$), were also recruited for participation. This group would only be asked to complete the standard self-efficacy instrument as outlined in the Data Sources and Collection section. The purpose for this control group was to compare any changes between students receiving the treatment and those who did not.

Recruitment. Students who were not enrolled in the C4 Scholar Program were recruited via email, since they were not in a cohort together. My hope was to recruit a similar sized control group to the treatment group. A list of these students was provided by a secretary at the university.

Potential participants received an email each day of the first week of class to provide an overview of the study, what their responsibilities as a participant would be, and what they would receive in return ($25 gift card to the university bookstore at the conclusion of the semester). I also reserved a classroom throughout the first week for 8 hours of walk-in hours (I detailed these in the email) for students to stop by to ask questions, review and sign their consent form, and/or complete their first survey. I also indicated that they could set up a time and location that was convenient for them. From past experience as an instructor at this university, I know that many students, especially incoming freshmen, experience difficulties with setting up their university email account and/or developing the habit to check their email on a daily basis. Because of this, I emailed the standard mathematics course instructors (all from the same course, but different sections) and asked if they would encourage their students to check their email and be on the lookout for my message. I heard back from many of the instructors and they were happy to do so. In addition, and similar to the extension given to the treatment group to gain more participants, I also extended the recruitment phase for the control group into the second week of classes. As well as sending reminder emails each day, I also visited a couple sections of the course to recruit in person. It was impossible for me to visit all sections because I taught a full course load myself. I chose to visit two sections in particular for a couple reasons. First, I did not teach or have any conflicts during their scheduled times. Second, I visited the two sections that held class during the scheduled walk-in hours that I hosted the previous week. After recruitment via email, in

39

person, and reminders from course instructors to check their email, I had 53 participants in the control group. It is important to note that I did not have any of these students in any other course that I taught.

Comparison of groups. Table 1 shows the demographic information of the 19 treatment participants and the 53 control participants who initially agreed to participate in the current study.

Table 1: Demographic Information for Treatment and Control Participants

		Treatment Group	**Control Group**
Gender	Female	10	35
	Male	9	18
Race	Black	3	16
	Hispanic	1	2
	White	15	35
Urban or Rural High School	Rural	14	27
	Urban	5	26
Average High School GPA	Out of 4.0	2.992	2.987
Average Total SAT Score	Between 400-1600	861.05	924.62
Average Math Section SAT Score	Between 200-800	444.74	444.06

A Chi-Square Test of Independence was used for gender, race, and whether participants' high school was urban or rural in order to see if the two groups were comparable in terms of their demographic information upon the start of the Fall 2019 semester. The results for (a) gender, χ^2 $(1, N = 72) = 1.07, p = .300$, (b) race, $\chi^2 (2, N = 72) = 1.51, p = .470$, and (c) urban/rural high school, $\chi^2 (1, N = 72) = 1.73, p = .188$ indicate that no differences were significant between the two groups. An Independent Samples T-test was conducted for High School GPA, Total SAT

scores, Math Section SAT scores, and *MSES* 1 scores. The results for those are discussed in the results section of this paper.

Description of Treatment

Participants in the treatment group of this study received the treatment described in this section. This treatment version of the developmental algebra course was initially designed to give students experiences of success with mathematics, focus on the content and skills they typically struggle with, and provide them with support as they persist and/or progress through the course. This course, as it was taught in Fall 2019, did not have any significant differences from how it was taught in Fall 2018 except for the one difference of a math autobiography journal, which will be detailed later. Sample assignments are provided in Appendix E. Betz and Hackett (1993) stated that if an intervention on self-efficacy were to be implemented, then they "recommend …[an] attempt to use all four sources of efficacy information in the design of math self-efficacy intervention programs" (p. 19). Even though this course was not designed around Bandura's (1977) theory on self-efficacy, strong connections to the four sources of self-efficacy were present and are described later. In order to enrich the description of the treatment, I will now present the guiding principles of the course, uniqueness of this course, and then how the four sources are linked to the course.

Guiding principles. The principles that act as the backbone of the design of this treatment are important to outline. These principles were held in mind by the researcher of the current study for the previous development of this specific math course for the C4 Scholar Program. They were formally written for the purpose of this dissertation. They include:

- Materials for the course were chosen with the following in mind: an activity's openness to accommodate a variation in student skill level (an open-ended problem or activity that

41

allows for more or less challenging exploration), opportunity for the emergence of success for these varying skill levels (challenge of solving the problem is within reach of all students with or without guidance from the instructor or peers), and the activity's ability to help students recognize value in the content and context of problems (student interests and real world contexts are incorporated).

- The importance of students' abilities to discuss mathematical ideas with their peers and instructor was stressed and fostered with collaborative and engaging activities, informal presentations of student work and ideas, and some student-led class activities (i.e. exam review, role playing as the expert for a concept, etc.). This combination of nontraditional instructional elements sets this course apart from others at the university.

- The instructor's expectations of students were realistically ambitious and reasonably paced for the duration of the semester. The course schedule was created by and adhered to by the instructor based on the integration of topics from the two-semester algebra sequence, which is further detailed in the next section. This schedule was successfully implemented with students in the past, as well as the current 2019-2020 cohort. Further, students were expected to put their best effort forth, and they received prompt feedback in return. The feedback that students received was a score on the assignment along with short comments about the success and/or struggle that was present for grading.

- A positive environment for active learning of mathematics was a priority for each class meeting. By active learning, this course encouraged participation from all students and supported them in doing so. A positive environment for this course was one that was supportive of students and their learning, open to everyone's questions and ideas, and a feeling of patience from all and towards all as open collaboration occurred.

The course as a treatment. There were many other sections of MATH 110 and 115 algebra courses at the university that similar incoming students would have taken before the development of the treatment course, some of which the researcher has taught. The treatment course, considered MATH 115 by the university, was different from the standard sequence, MATH 110 and 115 in several ways. The knowledge of these differences comes from the researcher's own experiences teaching this version and other versions of the course, conversations with other instructors, and conversations with the head of the mathematics department at the university. The main differences are presented below in four groupings: Content, Structure, Instructional Practice, and Assessments. These are outlined in more detail in Appendix F.

Content. The treatment course differed from other sections in terms of content. Included in content was the amount of attention topics in the course were given, and the coverage of topics in the treatment course was a subset of those covered in the other sections. Additionally, while application-based problems may conclude a lesson or unit in other sections of the course, the treatment section often began with a real-world problem to illustrate the content's utility and value right away. Other sections often began a lesson with necessary definitions and equations.

Structure. Differences in course structure also separated the treatment course from the other sections. For example, students must pay for a total of 7 credits to take both MATH 110 and 115. But, students in the treatment course (coded as MATH 115) only paid for 3 credits and still completed the sequence. In addition, the treatment course encouraged student attendance at two 75-minute workshops, as previously described. Other sections may offer a workshop, but it was required of all students up until a certain point in the semester (usually Exam 1, around 4 weeks) and then it was only required for students who did not meet a certain grade on an

assessment like an exam (i.e. 80%). These workshops were also led by a hired undergraduate student, not by the instructor.

Instructional practice. Instructional features of the treatment course included minimal lecture but with a supportive introduction of new topics, many opportunities for exploration and groupwork, and the instructor acted more as a facilitator than a lecturer. The instructor also moved freely around the room, instead of remaining at the front, to check-in with groups, talk to students, and answer questions in a more informal manner. The other sections, generally, utilize traditional instruction techniques like formal lecture, minimal exploration and groupwork opportunities, and the instructor acts a lecturer. This encourages fewer instances of interaction between students and the instructor or other students. Another difference is that I, the instructor of the treatment course, saw every submitted assignment, as compared to the other sections that use online homework. This allowed me to know where students were struggling, whereas the instructors of the other sections could obtain similar information if they clicked through the various assignments for each student.

Assessments. The treatment course and other sections of MATH 110 and 115 assess student learning of the course content in different ways. The treatment course used only paper and pencil assignments, which were graded and returned with handwritten feedback from the instructor. On these assignments, all work, including a final solution was required from students. If something was deemed incorrect, then an indication of where the error was made was noted by the instructor, but the correct answer was not provided so students could complete revisions. The other sections of the course, roughly 80% of them depending on the instructor, use online homework as a component of student grades. Online homework provides immediate feedback, but not from the instructor. Further, only answers are submitted and whole problems can be

44

marked incorrect. Students may not know where they made an error and this format does not "require" students to write out their work, although it is encouraged. The correct answer may be provided, and students are given the option to view an example or try a new problem.

Other differences included the treatment course's lesser weight on quizzes and tests, permission and encouragement to revise any graded item except the final exam, and the incorporation of Enhanced Learning Points (ELPs), which are also characteristic of all C4 courses. Students earned ELPs for doing nonrequired tasks and taking advantage of student-faculty interactions such as going to office hours, going to the workshop study sessions, going to the tutoring center, turning work in early, etc. These points were worth approximately 11% of their final grade and encouraged students to have these experiences which they should find to be positive and helpful. The other sections of similar courses weight formal assessments more heavily (i.e. other sections the current instructor taught were 75% of the grade), do not allow revisions, and do not utilize ELPs. These were the differences that separated the treatment course from other sections of the course.

Linking Bandura's sources to the treatment. This section will discuss the connections between features of the treatment course and Bandura's four sources of self-efficacy.

Mastery experiences. This is the most influential source (Bandura, 1977) and has the strongest link to the treatment. According to Bandura (1977), "successes raise mastery expectations; repeated failures lower them" (p. 195). The activities that students completed, along with appropriate scaffolding, helped students experience success. These experiences, which the treatment uniquely provided, lead students toward having mastery experiences, thus making the link to Bandura (1977). Examples of activities are in Appendix E. In particular, these activities were coupled with supports that could make experiences of mastery likely, if students

45

engaged with them. These supports include ELPs, as previously described, and optional revisions on any graded item except the final exam. For the revisions, students were not only earning points toward their grade, but it was the goal of revisions for students to carefully review their work, make corrections, and ask questions of their peers and instructor in order to earn a portion of the points that were not originally awarded. If students took advantage of these supports, which acted like training wheels, then they should see what they need to do to succeed in mathematics by having done so, i.e. going to office hours, reviewing their work, learning from mistakes, etc. Once students completed the course, they should have been able to see how they earned mastery and transfer these experiences to another math course or to their experiences with real-world mathematics. These "math-related performance accomplishments are important components of interventions intended to increase and sustain math self-efficacy over time." (Cordero et al, 2010, p. 363)

  Verbal persuasion. This source, which has weaker connections than mastery experiences in this treatment, was addressed directly from instructor-student interactions. This aligns with Bandura's (1977) definition of the source: "People are led, through suggestion, into believing they can cope successfully with what has overwhelmed them in the past" (p. 198). The instructor-student interactions varied greatly. Here is an example of an interaction from this semester, taken from the daily instructor journals. A student (who will later be identified as Participant 18) was very frustrated while trying to simplify exponential expressions. I told him that it is okay to ask questions, that I used to utilize office hours as a student, and that we just need to figure out what works for him. I reiterated that he was doing great and that I thought an issue he was having was that he was trying to simplify too much at one time. Then, I helped him

46

break down some of the problems and he did extremely well with the remaining examples, which seemed to improve how he appeared to be feeling at the time.

Verbal persuasion also occurred when I visited different groups during an interactive in-class activity. When I asked each group how their progress was coming along, some groups responded positively while others expressed their struggles. If the latter was the case, I asked for and looked at the work they completed. If I had to point out an error for their consideration, I also explained the connection to the correct answer so they did not feel discouraged and they could see their progress. No matter if the group successfully completed an activity or struggled, I always engaged them in a positive and productive conversation. Another example of verbal persuasion were my meetings, often during office hours, with individual students. During these meetings, I discussed individual progress, any missing assignments, answered questions, and discussed what the student felt s/he was struggling with. We often came up with a plan of action together. Encouragement was an integral part of these meetings and I always made sure to start and end with what the student was doing well with so that s/he did not leave feeling defeated. It is important to note that "verbal persuasion has limitations, but can be useful when used in conjunction with opportunities for performance practice" (Hodges, 2008, p. 266).

Emotional arousal. This source is important because "high arousal usually debilitates performance, [and] individuals are more likely to expect success when they" are more at ease (Bandura, 1977, p. 198). But emotions may be either positive or negative. When students walked into the classroom each day, they received positive affirmations of their presence which hopefully put them at ease and lowered their emotional arousal. I always greeted them as they entered the classroom and I asked them about their outside lives (i.e. "How'd your job interview go?" or "How was your dinner with friends last night?"). I did not want them feeling stressed or

anxious just by walking into math class. Additionally, activities, like those in Appendix E, provided positive feelings of safety and could lessen negative feelings of arousal because students were able to check their answers and the multiple entry points catered to varying skill levels. However, these seemingly safe activities also encouraged discussion and creative thinking. By creating a warm environment and engaging class activities, a positive effect on the emotions they associated with mathematics and their self-efficacy toward the subject was hopefully experienced and emotional arousal was hopefully reduced.

Vicarious experiences. The final source has proved to be difficult to measure in recent research (Usher & Pajares, 2008). Vicarious experiences are had via "inferences from social comparison" which could come from observation and/or learning about a similar person's experiences (Bandura, 1977, p. 197). Opportunities for vicarious experiences were provided by planning frequent in-class group work and encouraging collaboration outside of the classroom. In-class groups were arranged by the instructor and stronger students were often paired with those that may struggle. Groups changed every two to three weeks. Creating groups in this manner should have supported vicarious experiences between the participants. Additionally, Betz and Hackett (1993) encourage the use of role models for this source of self-efficacy. The student data collector who interviewed the participants was in the C4 Scholar Program two years ago and experienced much success as a result, even recently earning a scholarship. Introducing him at the beginning of the semester and detailing his success could have also been a source of vicarious experience, because the students may have viewed him as a role model. Participants were told they would have future interactions with him during the interviews.

Data Sources and Collection

I utilized mixed-methods data collection strategies. The main strategies included a

standard measure of self-efficacy (collected three times), Journals (collected five times), and student Interviews (collected two times). In addition, the grades that treatment participants earned on major assessments and instructor journaling were collected to address research question four because these could help to further explain a student's self-efficacy (Pajares, 1996). Each measure was collected from all participating treatment group students during Fall 2019. Only the standard measure of self-efficacy was collected twice from the control group participants.

Mathematics Self-Efficacy Scale. Research participants in both groups completed the *Mathematics Self-Efficacy Scale (MSES)*. I used this instrument because it is commonly used and validated in self-efficacy research. Developed by Betz and Hackett (1993), the *MSES* is composed of 34 questions, each with a 10-point Likert response scale. It takes approximately 15 minutes to complete and is targeted for use with college student populations. The first part (Everyday Math Tasks) asks participants questions about their beliefs in their ability to solve problems involving mathematics in everyday situations and in general (i.e. balancing a checkbook or adding two large numbers in one's head). The second part (Math Courses) asks participants about their beliefs in their ability to achieve success in various levels of college courses where mathematics is required. The *MSES* produces a total score between 0 "No Confidence" to 9 "Complete Confidence;" it does not separately measure Bandura's four sources of self-efficacy in mathematics.

Treatment group. Research participants in the treatment mathematics course completed the *MSES* three times – at the beginning (week 1 or 2), middle (week 8), and end of the semester (week 15). Before each administration, I reviewed the directions of the survey on the document camera, reminded them that I would not see their responses until final grades have been

submitted, encouraged them to consider and reflect on their experiences in the current mathematics course (except for the first administration), and told them that they could reschedule the *MSES* for the next day if they felt as though their responses would be an inaccurate representation of their experiences, i.e. if they were having a very bad day for any reason. The student data collector was present for each administration. To avoid undue influence, I stepped outside the classroom as students took the survey. The student data collector distributed paper copies of the survey. Once the surveys were completed, she placed all surveys, which each had a cover sheet for confidentiality of their responses, in a sealed envelope. The student data collector also recorded each submission on the data tracker sheet so that she could follow up with any participants who were absent. Completed surveys from the three administrations were stored in separate sealed envelopes and placed in a locked cabinet in my locked office.

   Control group. Research participants who took the standard two-semester mathematics sequence completed the *MSES* twice – at the beginning (week 1 or 2) and end of the semester (week 15). To complete the first administration, I emailed the participants in the standard mathematics course (as described in the Participants section) to offer walk-in hours or encouraged them to set up a time to meet with me at their convenience. The first survey was completed alongside the consent form. Because I was not their instructor, I did not expect that my presence would influence their responses in any appreciable way. I emailed participants again at the end of the semester in advance of completing the second administration. Before each student took the survey, I reviewed the directions for both administrations and for only the second administration encouraged them to reflect on their experiences related to the current math course. Completed surveys from both administrations were stored in separate sealed envelopes and placed in a locked cabinet in my locked office.

Journals. Five Journals, which asked students to describe how they currently felt about mathematics, were collected. These Journals asked all students in the treatment mathematics class to free-write about their experience with the recent unit in the course (how they felt about the material, successes or failures they experienced, how they felt about a recent exam, etc.). At the conclusion of each unit, a Journal was administered. Before each administration of the Journal, I reviewed the directions of the Journal on the document camera, reminded them that I would not see their responses until final grades have been submitted, and encouraged them to consider and reflect on their experiences in the current mathematics unit (except for the first administration since it was a mathematics autobiography). Students were encouraged to write whatever came to mind, but a standard menu of prompts were provided in case some were unsure about what to write. See Appendix G for Journal prompts 1 – 5. The student data collector was present for each administration. To avoid undue influence, I stepped outside the classroom as students wrote their responses. The student data collector distributed paper copies of the Journal. Once they were completed, all Journals were placed in a sealed envelope. The student data collector also recorded each submission on the data tracker sheet so that she could follow up with any participants who were absent. Completed Journals from the five administrations were stored in separate sealed envelopes and placed in a locked cabinet in my locked office. Participants were given approximately 20 minutes of class time, at the end of the last class of each unit. If anyone needed more time, they would have been accommodated without interruption, either in the classroom or the nearby study room. This, however, never occurred.

- The first Journal (week 1) was a math autobiography.
- The second Journal (week 5) was on unit 1.

- The third Journal (week 10) was on unit 2.

- The fourth Journal (week 14) was on unit 3.

- The fifth Journal (week 15) was a reflection on the semester.

Interviews. Audio-recorded Interviews were carried out at the beginning and end of the semester (Interviews 1a/b and 2a/b) by a second student data collector. Some participants in the treatment group were called upon for a follow-up Interview after 1a or 2a for any necessary clarifications (Interviews 1b and 2b). All participants in the treatment group completed the 1a and 2a Interviews, but not all had the 1b and 2b Interviews. Five participants had the 1b Interview and one participant had the 2b Interview. Interviews asked participants to describe their background/experiences, narrate the changes that they had been experiencing (if any), and were used, alongside data analysis, to further understand the level of self-efficacy a student has for mathematics from his/her point of view. Questions for Interviews 1a and 2a are similar, but the student was asked to focus on their past experiences (high school and earlier) in 1a and on their experiences during the Fall 2019 semester in 2a. Both Interview protocols were written for use in the current study and were informed by Bandura's (1977) four sources of self-efficacy. Questions for Interviews 1b and 2b aim to clarify any responses from 1a and 2a as the researcher deemed necessary. See Appendix H for Interview protocols.

To mitigate any influence from me (the instructor), a student data collector carried out the Interviews. I met with the student data collector before, during, and after each set of Interviews. Details of preparing the student data collector are provided in Appendix I. Meeting before the Interviews served the purpose of reviewing the Interview protocol and clarifying the questions. Meeting during the first set of Interviews served the purpose of answering any further questions that the student data collector had. Details of how this meeting benefited the student data

collector and how it affected the overall quality of the Interviews are in Appendix I. Meeting after the Interviews served the purpose of identifying any students that needed to be called upon to participate in a 1b or 2b Interview. Before this meeting, I listened to audio recordings of each Interview and took notes on anything that should be clarified in a follow-up Interview. Once the Interviews were complete, the audio recordings were transcribed.

Pilot interviews. The Interview 1a protocol was piloted in August 2019 with student volunteers. Three students who participated in the C4 Scholar Program in the past were the volunteers. These students were contacted via email and volunteers were taken. My goal was to have 3-5 volunteers. These pilot interviews served the purpose of developing follow-up questions and fixing any questions in the interview protocol. The pilot interviews occurred with the student data collector. The student data collector observed the entire first pilot interview. The student data collector observed half of the second pilot interview and then asked the remaining questions. The student data collector led the entire third pilot interview. After each pilot interview was complete, we asked the volunteer about their experience during the interview. Then, after the volunteer left the conversation, the student data collector and I discussed the questions along with the volunteers' responses, assessed the flow of the interview, and made notes throughout the protocol of statements the student data collector could make to help the interview process run smoothly (i.e. clarifying directions, providing an introduction to a set of questions, etc.).

Plan for student withdrawal. Students in the treatment mathematics class could have withdrawn from the course or the university at any point in the semester. However, all participants in the treatment course finished the Fall 2019 semester. If this would have happened, the student (only those from the treatment mathematics class) would have been asked, if willing,

to participate in one last round of data collection for all measures upon their withdrawal, if possible. That means, they would take *MSES* one last time so that a final measure of self-efficacy was recorded, they would complete the Journal that was closest to their departure if it had not already been written, and they would complete the Interview that was closest to their departure if it had not already occurred. The measures of self-efficacy would have been from the start of the semester to the time of withdrawal for the individual student.

Student grades. Student grades on four major exams were collected in order to compare to the results of the other data sources and answer the fourth research question. These include Exam 1 (over Unit 1, which covers Order of Operations, Fractions, Ratios and Proportions, and Percentages), Exam 2 (over Unit 2, which covers Linear Equations, Linear Systems, Exponents, and Exponential Equations), Exam 3 (over Unit 3, which covers Logarithmic Equations, Factoring and Quadratic Formula, and Quadratic Equations), and a cumulative Final Exam. Thus, another form of data collection was the instructor's gradebook. This was electronic in form and could filter individual students or compare the whole group. Grades on each of the exams were recorded as *points earned* out of 100 points possible. Final grades were reported as letters.

Instructor journals. The researcher's observation of events that may have provoked changes in student self-efficacy were attempted to be documented in daily journals. More specifically, these journals discussed any observed instances, by the instructor, that a student may have experienced influence on their self-efficacy, whether it be positive or negative. Prior to the Fall 2019 semester, the researcher recognized that this would be challenging because instances would be circumstantial, they may occur while I was with another student, they may have occurred outside of the classroom, and/or they may have been difficult to see since a student's self-beliefs are not a physical characteristic. The researcher made brief notes during

class and expanded on them immediately after. There was one daily instructor journal per class period except for each Wednesday. On Wednesdays, the students had a regular class period and then a math workshop period which followed 40 minutes later. On these days, two entries were made: (a) one journal for the regular class period was written during the 40 minutes between the two sessions and (b) a second journal for the workshop was written following the conclusion of that period. This was done so that events during the two periods were not intermixed. These were written in a running Word Document and kept on the instructor's computer which was password protected. In summary, the instructor journals attempted to document observable instances that may have led to changes in students' self-efficacy and notes about the class period in general including any instructional decisions that may have caused change to the anticipated plan.

Timeline of events. A timeline of when major assessments and data collection occurred during the Fall 2019 semester is observed in Figure 1.

Figure 1: Timeline of Events in Fall 2019

Start of Semester Fall 2019

Week 1
• Recruitment
• MSES 1
• Journal 1

Week 2
• Recruitment
• MSES 1
• Journal 1
• Interview 1a
• Interview 1a

Week 3

Week 4

Week 5: Exam 1, Journal 2 Interview 1b

Week 6

Week 7
• Start of Drama

Week 8
• MSES 2

Week 9

Week 10: Exam 2, Journal 3

Week 11

Week 12

Week 13

Week 14: Exam 3, Journal 4, Interview 2a

Week 15
• MSES 3
• Journal 5
• Interview 2a/2b

Final Exam and End of Semester Fall 2019

Data Analysis

As a reminder, here are the research questions that guided the current study:

1. What, if any, changes in mathematics self-efficacy occur for at-risk college students, as measured with a standard instrument, when the instructional treatment is implemented?

2. How do these results compare to students of similar placement, but who do not receive the treatment, as measured with the same instrument?

3. According to students' journals and interviews from the treatment group, have changes in their level of self-efficacy occurred? And, what do they think are the reasons for these changes or absence of changes?

4. What consistencies and/or or inconsistencies exist between the results of each of the collected data sources: (a) the standard instrument (*MSES*), the students' perspectives through (b) journals, (c) interviews, (d) students' course assessments, and (e) the instructor's perspective, which was motivated by documented instances of possible influence on participants' self-efficacy throughout the semester in daily instructor journals?

Table 2 shows how each of the data collection measures corresponds to a research question(s).

Table 2: Relation of Measures to Research Questions

Research Question	Measure(s)	Explanation
1	*MSES* Questionnaire (Weeks 1, 8, 15)	This questionnaire measured whether changes in self-efficacy occurred (increase, decrease, no change) across the semester.
2	*MSES* Questionnaire (Weeks 1, 15)	This allowed for a comparison between self-efficacy levels of the treatment and control groups.
3	Student Journals (Weeks 1, 5, 10, 14, 15) and Student Interviews (Weeks 2-3, 14)	These helped the researcher to understand, via analysis, whether students' self-reports indicate changes in self-efficacy, if any.
4	*MSES* Questionnaire (Weeks 1, 8, 15), Student Journals (Weeks 1, 5, 10, 14, 15), Student Interviews (Weeks 2-3, 14), Course Assessments, and Instructor Journals	These offered a comparison between the results of all five data sources to see what level of agreement occurred. A discussion of the analysis for each pairing will be presented.

Addressing research question 1. This question, which required quantitative analysis, asks if changes of any type occurred in treatment participants' self-efficacy as measured by the *MSES* instrument. According to the *MSES* manual, the sum of numerical selections (0-9 for each question) divided by the number of questions (34) yields the respondent's self-efficacy score. Higher scores indicate greater self-efficacy. The *MSES* manual (Betz & Hackett, 1993) contains the following confidence scale:

- No Confidence at all: score of 0

- Very little Confidence: score of 1, 2, 3

- Some Confidence: score of 4, 5

- Much Confidence: score of 6, 7

- Complete Confidence: score of 8, 9

It is important to note that it was rare for any students' scores on the *MSES* to be whole numbers. For the purpose of this study, I referred to the first score (during Week 1) as the student's base score—the student's level of mathematics self-efficacy when s/he entered the course. In contrast to their base scores, students' subsequent scores may have increased, decreased, or did not change. So, there were nine possible outcomes for the three *MSES* scores for each student:

A. Score 1: Base, Score 2: Increase, Score 3: Increase

B. Score 1: Base, Score 2: Increase, Score 3: Decrease

C. Score 1: Base, Score 2: Decrease, Score 3: Increase

D. Score 1: Base, Score 2: Decrease, Score 3: Decrease

E. Score 1: Base, Score 2: No Change, Score 3: Increase

F. Score 1: Base, Score 2: No Change, Score 3: Decrease

G. Score 1: Base, Score 2: Increase, Score 3: No Change

H. Score 1: Base, Score 2: Decrease, Score 3: No Change

I. Score 1: Base, Score 2: No Change, Score 3: No Change

Participants' scores on the two subscales and the entire survey (34 questions) were considered. To measure the changes in the participants' *MSES* scores, the program SPSS was used to indicate the level of significance of any changes, if there were any. In particular, a Repeated Measures ANOVA was used to test for differences in mean total *MSES* scores for the treatment participants across time. I have worked with SPSS in the past, but to reorient myself with the program, I worked with a colleague of one of my guidance committee members. She was also able to discuss which tests would help to best tell the story of the participants' experiences, in addition to the results that were found from implementing these tests.

Addressing research question 2. This question, which also required quantitative analysis, focuses on comparing the results of the standard instrument for the treatment and control groups. Recall unlike the participants in the treatment course, the control group participants were not all enrolled in the same section of MATH 110. A comparison of control participants who did and who did not complete the study was made using a Chi-Square Test of Independence and an Independent Samples Test (T-test). A Repeated Measures ANOVA was used to see if significant differences in the control group's *MSES* scores occurred across time, as was done for the treatment group. An Independent Samples Test (T-test) was used to compare the two groups' high school GPAs, Total SAT scores, Math Section SAT scores, results of the first *MSES* administration, and results of the two subscales of the *MSES*. A Chi-Square Test of Independence was used to compare the two groups on the bases of gender, race, and urban/rural high school. To consider the final results of the two groups, a Repeated Measures ANOVA was used to compare the changes made by both groups across time for total *MSES* score and the two

subscales to determine if a time-by-treatment interaction exists. The effect size of the study was considered and will be discussed.

Addressing research question 3. To answer the third research question, which is qualitative in nature, relevant data from participant Journals and Interviews were identified and coded. This question focuses on whether the treatment participants' Journals and Interviews revealed whether a change was occurring or did occur with their self-efficacy, answered by data analysis, and how the participants made sense of these perceived changes or absence of change. To begin answering this question, I searched for statements from the treatment students in their written Journals and Interview transcripts that appeared to indicate mathematics self-efficacy and/or changes in it. Statements such as, "I think I'm going to do really good on the next test," "I've never earned an A on a math quiz before this class," or "I'm still no good at math" were considered as statements about mathematics self-efficacy. Upon development of the codebook, which is discussed in the following section, these statements were coded as Positive, Negative, or Variable and then coded further into Past, Present, or Future. The reasons that participants provided statements such as these were also identified and coded. Table 3 illustrates the codes, definitions, and examples from the data that were used for the analysis. The definition of each code is based on the definition of self-efficacy that was provided in Chapter 2 by Bandura (1977). Recall, self-efficacy is the "beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments" (Bandura, 1997, p. 3). The discussion following the table describes the development of the codebook and the coding process.

Table 3: Codes, Definitions, and Examples for Research Question 3

SELF-EFFICACY CODE	DEFINITION	EXAMPLE
1: Positive (+)	Statement that is indicative of a positive belief about being successful with mathematics. The level of positivity can vary.	"The way I felt after the Exam 2 score was very confident in myself" (Journal 4, p. 9).
(+) Past	Positive statement (definition above) that speaks of the past (before college)	"I did really well in the subject, it's actually one of my favorites" (Journal 1, p. 31 – when talking about middle/high school)
(+) Present	Positive statement (definition above) that speaks of the present (current semester). These statements can be further broken down into shorter-term feelings (<1 month) or longer-term feelings (>1 month).	"I never had a B+ or A- in math before and I'm proud of myself for that" (Journal 5, p. 3 – longer-term feeling)
(+) Future	Positive statement (definition above) that speaks of the future (after current semester)	"I feel really different, I can go into other courses feeling like I actually know what I am doing without any help" (Journal 5, p. 3)
1a: Reason for positive	When a statement for code 1 is identified, a code of 1a will identify reasoning if provided	"The result was what I expected because I put the time and effort in studying" (Journal 4, p. 9). Accompanies first positive code example.
2: Negative (-)	Statement that is indicative of a negative belief about being successful with mathematics. The level of negativity can vary.	"I have always struggled with math" (Journal 1, p. 7). "I don't usually perform very well in math classes…I was just bad at it" (Journal 1, p. 23). *Two clauses here, both receive a negative code.
(-) Past	Negative statement (definition above) that speaks of the past (before college)	"I have always struggled with math. It has been my toughest subject" (Journal 1, p. 7)
(-) Present	Negative statement (definition above) that speaks of the present (current semester). These statements can be further broken down into shorter-term feelings (<1 month) or longer-term feelings (>1 month).	"I felt terrible about my…exam 3 [score]" (Journal 5, p. 7 – shorter-term feeling)
(-) Future	Negative statement (definition above) that speaks of the future (after current semester)	"I also am worried that I'm gonna fail my class because of this one exam" (Journal 4, p. 29)
2a: Reason for negative	When a statement for code 2 is identified, a code of 2a will identify reasoning if provided	"I can try to solve problems, but I somehow always end up messing up one part, which screws up the rest" (Journal 1, p. 7). Accompanies first negative example.
3: Variable	Statement that is indicative of a mixture of both positive and negative beliefs about being successful with mathematics. The levels of positivity and negativity can vary.	"My experience this semester was like a roller coaster" (Journal 5, p. 31). "My experience this year in math has been all over the place" (Journal 5, p. 33).

Developing the codes. Originally, I applied codes of Positive or Negative to applicable

statements to the data based on a simple version of Table 3. An exploration of whether a

"neutral" code should exist was conducted. It was determined that there should not be a "neutral" code because statements that did not fall under Positive or Negative categories were more factual or report-like in nature and did not indicate how the participant was feeling about him/herself with mathematics. The process that led to this decision is detailed in the next paragraph.

I opened the data set provided by Journal 5. I read through the first half of the submissions from half of the participants. I considered statements that were not highlighted to determine if I would have assigned those a neutral code if one existed at the time. I determined that the statements that were not highlighted seemed factual. For example, "My exam 3 score was the lowest out of all my exams" (Journal 5, p. 3) or "I really like that we can get exact answers to most math problems" (Journal 5, p. 9). If there was a comment that indicated a positive or negative feeling that preceded or followed, then that preceding or follow-up statement was coded. Therefore, statements without a positive or negative feeling were not coded. I did not believe that statements like these would affect the rating of the Journal. However, just to be sure, I opened Journal 3 and started reviewing those responses. I made the same conclusion about these. For example, Journal 3 statements that had feeling to them were coded and statements that did not (more factual) were not coded – here is an example: "The assignments in this unit have been complex but also very educating" (Journal 3, p. 19). I did not feel this statement told me much about the participant's self-efficacy or how these assignments affected the participant. Other statements in this Journal did indicate positive and/or negative feelings about his/her self-efficacy and so those were coded as described in the codebook.

Following this development, the researcher met with MLRG at Michigan State University to receive feedback on the codebook in February 2020. Issues surrounding variable statements (those that indicated positive and negative feelings), how students defined success, and the

length of time that reported feelings influenced students' self-efficacy were discussed. I addressed each of these and after returning to the data, decided it was appropriate to expand the codebook, which is what is reflected in Table 3. First, Positive and Negative codes would be further categorized by Past, Present, and Future statements depending on if participants spoke about an experience before college, during the current course/semester, or looked to the future after the course. In addition, Present codes would be further broken down into how long the feeling could affect the participant's self-efficacy toward math (i.e. shorter-term feeling < 1 month or longer-term feeling > 1 month). One month was chosen as the division between shorter-term and longer-term because that was roughly the length of each unit in the course and was the amount of time that separates most of the Journals. Second, on the same level as Positive and Negative codes, a Variable code was added in order to address the statements participants made about their experience being a mixture of positive and negative. Examples for each of these are included in Table 3.

   Levels of coding. There were three levels of coding that were applied to the Journals and Interviews. First, clauses and small statements were rated with the codes Positive, Negative, or Variable, in relation to self-efficacy, as shown in Table 3. A single statement could be coded into multiple clauses. A participant may have included two or more positive clauses in a single statement, two or more negative clauses in a single statement, or a combination of positive and negative clauses. Second, in addition to assigned Positive or Negative codes, a time frame was also assigned. These include Past, Present, or Future. Past refers to before college, Present refers to the current course/semester, and Future refers to after the current course/semester. For statements that were coded Present, they could be further coded into how long the positive/negative feeling may impact the participant if it were obvious (i.e. shorter-term feeling,

<1 month or longer-term feeling, >1 month). Third, ratings for Journals and Interviews were given. These ratings range from "Very Negative" to "Very Positive" self-efficacy. This process is detailed in the next section.

Assigning a rating. After Journals and Interviews were coded using Table 3, a count of the number of Positive and Negative self-efficacy clauses/statements and reasons was determined. This count was used to assign a rating to the participant's Journal or Interview based on Table 4. A similar scale to the categories given by the *MSES* confidence scale was developed for this study. Please note for the percentages below, the total number of coded statements make up 100%. To illustrate this, let us consider a participant's Journal with 3 negative statements and 5 positive statements. There are 8 statements total, leaning in the positive direction. To determine a rating, I would calculate $5 \div 8 = 62.5\%$, which is a Somewhat Positive rating. This method was applied to determine a rating for each Journal and Interview for every participant in the treatment group. It is important to indicate that advice from and close work with my advisors resulted in the choice to construct and analyze the qualitative data from Journals and Interviews into ordinal scales. The justification for this choice was to have a relatively principled way to compare and discuss the results of data analysis for multiple sources of data in Research Question 4. However, even though quotes will be used throughout the response to Research Question 3 to support the ratings from the ordinal scales, the overall descriptive richness of the data lessens. This richness might have been preserved from choosing a different direction for data analysis, such as individual case studies.

positive statements + negative statements = 100%

Table 4: Rating Scale of Journals and Interviews

Rating Category	Description	Percentages of Statements
Very Negative Self-efficacy	all negative, no positive statements OR overwhelmingly negative	85%-100% of the statements are negative
Moderately Negative Self-efficacy	many more negative than positive statements	75%-84% of the statements are negative
Somewhat Negative Self-efficacy	a few more negative than positive statements	61%-74% of the statements are negative
Neutral	fairly equal negative and positive statements	50%-60% of the statements are positive and/or 50%-60% are negative
Somewhat Positive Self-efficacy	a few more positive than negative statements	61%-74% of the statements are positive
Moderately Positive Self-efficacy	many more positive than negative statements	75%-84% of the statements are positive
Very Positive Self-efficacy	all positive, no negative statements OR overwhelmingly positive	85%-100% of the statements are positive

Further, I paired colors with the ratings in order to clearly illustrate any increases, decreases, or no change in self-efficacy experienced by participants in a visual manner, see Figure 2.

Figure 2: Rating Scale of Journals and Interviews with Colors

Very Negative	Moderately Negative	Somewhat Negative	Neutral	Somewhat Positive	Moderately Positive	Very Positive

Variable statements were not included in this count because they could not be categorized as just positive or negative. However, the number of variable statements were used to indicate whether a participant made note of any changes or variation happening during their experience in the treatment course. A comparison across time was made. A table of the number of variable statements for each participant in each data piece will be provided and discussed in Chapter 4.

After coding for the Journals and Interviews was complete, a second stage of analysis took place for research question three. This stage considered the ratings for Journals 1, 5 and

Interviews 1a/b, 2a/b and made note if an increase, decrease, or no change in self-efficacy occurred, according to the analysis of participant responses and Table 4. To clarify, the ratings for Journal 1 and Interview 1a/b (similarly for *MSES* 1) were considered base self-efficacy ratings for each participant. The ratings from Journal 5 and Interview 2a/b were then compared to the ratings of Journal 1 and Interview 1a/b, respectively. An overall indication of increase, decrease, or no change in self-efficacy for each participant was given for Journals and Interviews. The participants were, then, categorized based on the combination of increase, decrease, or no change that their Journals and Interviews exhibited (e.g. increase on one measure and decrease on another measure, etc.). Each category of participants, with support from participant statements, will be discussed in Chapter 4. This method (coding and analysis of Interviews and Journals) was systematic in nature. The researcher followed the same procedure for each participant.

The preceding analysis answers the first part of research question three. To answer the second part, which focuses on the reasons that participants reported for any observed changes in self-efficacy, as identified by analysis, I identified the topics that participants most frequently wrote or spoke about. This was done on an individual and group level. First, for individuals, I read through all Journals and Interviews and identified topics that participants discussed that appeared to affect them. It is important to indicate that the topics emerged from the data, I did not begin looking at the data with preconceived ideas or a list of any sort. For example, one topic that was commonly discussed by participants that emerged from the data were experiences of success and/or struggle that seemed to affect how they felt about mathematics. For each participant, I tallied the number of instances that they appeared to report something (i.e. an element of the course or program) that had an effect on them. Second, for the group, I combined

the tallies for each of the Journals and Interviews to determine which topics were most reported by the group. Upon conclusion of this focus of analysis, six topics were identified as being reported on most often. These will be identified and discussed, along with examples, in Chapter 4. In addition, if a topic was an intentional aspect of the treatment course or C4 Scholar Program, then a discussion about how it is specific to the course and program will be discussed in Chapter 4 as well. The analysis completed for individuals can help to further understand how an individual's self-efficacy may have changed throughout the semester. The analysis completed for the group can help to understand what the common factors are that appear to affect self-efficacy.

Addressing research question 4. To answer the fourth research question, which asks about the consistencies and/or inconsistences that exist between collected data sources, a comparison of data analysis results for each data source was considered. When the results of this research question are presented in Chapter 4, the results of the standard instrument and participants' perspectives through Journals and Interviews will have already been introduced. The results of the two data sources that remain, treatment participants' grades on course assessments and the Instructor's Perspective, will be first reported and then a comparison between each of the results of the five data sources will be discussed. In this section of data analysis, I will first discuss the final two data sources and then I will discuss the method I used to assess agreement between the five data sources.

Participants' course assessments. The grades that participants earned on three distinct unit exams and one final exam were considered for the current study. Initially, their quiz scores were also considered. However, because the experience with quizzes 1 and 2 were non-traditional (an overwhelming number of questions on Quiz 1 and a C4 Scholar Program related learning experience on Quiz 2) and the exams seemed to carry much more weight for the

participants (point-wise and emotionally), the decision to exclude the quizzes was made. The four exam scores were each out of 100 points. Exam 1 occurred the day after Journal 2, Exam 2 occurred the day after Journal 3, Exam 3 occurred the day after Journal 4, and the final exam occurred five days after Journal 5. In addition to presenting the exam scores, a spaghetti plot will be presented to illustrate the scores in a graphical manner.

Instructor perspective. The Instructor Perspective was formed from the documented experiences that I journaled about after every class period throughout the Fall 2019 semester. These journals attempted to capture any instance that I believed could affect the participants' self-efficacy toward mathematics. The Instructor Perspective is comprised of three ratings that span three time periods throughout the semester. These time periods corresponded with the content units in the course and were assigned at Week 5 (which consisted of weeks 1-5 and Unit 1), Week 10 (which consisted of weeks 6-10 and Unit 2), and Week 16 (which consisted of weeks 11-16 and Unit 3). I did not give students an initial (or start of the semester) rating because I did not know them when they entered my classroom. I knew all their names and faces by the end of Week 1.

I considered assigning weekly ratings for each participant, and attempted this, but the observations of participants' experiences and self-efficacy did not vary enough week-to-week to justify a weekly (x16) rating as compared to a unit rating. I did not feel that anything was gained from having more ratings because students' self-efficacy and experience seemed to change more each unit rather than within units or weekly. Additionally, students appeared to frequently think about the course as units and would refer to Unit 1, 2, or 3 often, rather than speaking about individual weeks.

To define the Instructor Perspective, I followed the steps below on two separate occasions to be as close to accuracy in these ratings as possible:

1. I scanned my instructor journals from the semester and searched for each of the 19 participants' names.

2. For each timeframe (i.e. Week 5 consisted of weeks 1-5, Week 10 consisted of weeks 6-10, and Week 16 consisted of weeks 11-16), I considered all entries for each student and coded them as positive or negative entries.

   o *Example of positive entry*: "Participant 5 asked me [after class] about the deer proportion problem – and was SO EXCITED to learn that she did it correctly. She sort of jumped up and down and said she was feeling more confident." (Week 4)

   o *Example of negative entry*: "Participant 3 said, 'I just want to get in my car and drive away' at the beginning of class [before Quiz 1]." (Week 3)

3. After I identified entries in the instructor journal that should receive positive and negative codes, a count of the number of positive and negative entries was determined for each time period. Then, a rating was determined, as described below.

I assigned ratings based on the categories in Figure 3 and descriptions in Table 5 below. Please note that these ratings were constructed from the scale for Journals and Interviews: ranging from Very Negative to Very Positive in order to show comparison. This was done for each of the 19 students across three timeframes. Please note for the percentages below, the total number of coded entries in each time frame make up 100%. To illustrate this, let us consider a participant who had 2 negative entries and 4 positive entries for a single time period. There are 6 statements total, leaning in the positive direction. To determine a rating, I would calculate 4 ÷ 6 = 66.7%, which is a Mostly Positive rating. It is, again, important to indicate that advice from

and close work with my advisors resulted in the choice to construct and analyze the qualitative

data into ordinal scales, this time for instructor journaling. The justification for this choice was to

have a relatively principled way to compare and discuss the results of data analysis for multiple

sources of data in Research Question 4. However, this method of data analysis can lessen the

data's descriptive richness. This richness might have been preserved from choosing a different

direction for data analysis, such as individual case studies.

$$\text{positive entries} + \text{negative entries} = 100\%$$

Table 5: Rating Scale of Instructor Perspective

Rating Category	Description	Percentages of Statements
Very Negative Self-efficacy	all negative, no positive entries OR overwhelmingly negative	85%-100% of the entries are negative
Mostly Negative Self-efficacy	many more negative than positive entries	65%-84% of the entries are negative
Mixed Self-efficacy	fairly equal negative and positive entries, can have a note about if the participant is slightly more positive or negative within mixed	50%-64% of the entries are positive and/or 50%-64% are negative\n\n*Note: 60-64% in either direction receive the Mixed-Positive or Mixed-Negative notation
Mostly Positive Self-efficacy	many more positive than negative entries	65%-84% of the entries are positive
Very Positive Self-efficacy	all positive, no negative entries OR overwhelmingly positive	85%-100% of the entries are positive

A visual representation of the scale for the Instructor Perspective is given in Figure 3.

Figure 3: Self-Efficacy Scale Used for the Instructor Perspective
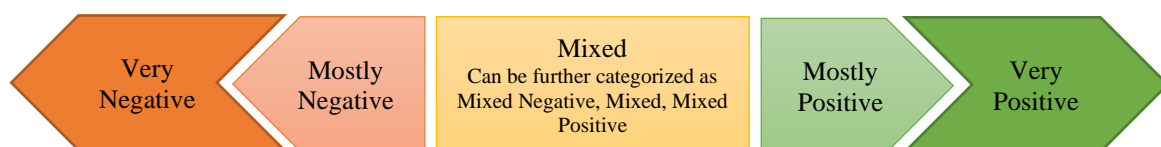
Further, the descriptions of when specific ratings apply to each time period are as follows:

- **Very Negative** – Participant <u>only or primarily had recorded experiences of struggle</u> in the course during the specified time period (i.e. only earned poor grades). Participant was observed to <u>primarily experience struggle</u> (i.e. showed frustration, made negative comments, negative body language and composure in class). Participant made <u>primarily negative comments</u> about him/herself in relation to mathematics (i.e. "I feel stupid"). Very seldom or no positive experiences or sentiments were observed.

- **Mostly Negative** – Participant <u>mostly had recorded experiences of struggle</u> in the course during the specified time period (i.e. mostly earned poor grades). Participant was observed to <u>mostly experience struggle</u> (i.e. mostly showed frustration, mostly made negative comments, mostly had negative body language and composure in class). Participant made <u>mostly negative comments</u> about him/herself in relation to mathematics (i.e. "I feel stupid").

- **Mixed**

  o **Mixed Negative** – Participant <u>had slightly more recorded experiences of struggle</u> in the course during the specified time period (i.e. generally earned poor grades). Participant was observed <u>to experience struggle slightly more</u> (i.e. generally showed frustration, generally made negative comments, had slightly more negative body language and composure in class). Participant made <u>negative comments slightly more</u> about him/herself in relation to mathematics (i.e. "I feel stupid").

  o **Mixed** – Participant had <u>equal recorded experiences of success and struggle</u> in the course during the specified time period. Participant was observed to <u>experience</u>

71

success and struggle equally (i.e. showed equal frustration and excitement, made equal positive and negative comments, showed equal positive and negative body language and composure in class). Participant made equally positive and negative comments about him/herself in relation to mathematics.

- o **Mixed Positive** – Participant had slightly more recorded experiences of success in the course during the specified time period (i.e. generally earned quality grades, indicative of mastery). Participant was observed to experience success slightly more (i.e. generally showed excitement, generally made positive comments, generally had slightly more positive body language and composure in class). Participant made positive comments slightly more about him/herself in relation to mathematics (i.e. "This is easy." or "I know how to do this!").

- **Mostly Positive** – Participant mostly had recorded experiences of success in the course during the specified time period (i.e. mostly earned quality grades, indicative of mastery). Participant was observed to mostly experience success (i.e. mostly showed excitement, mostly made positive comments, mostly had positive body language and composure in class). Participant made mostly positive comments about him/herself in relation to mathematics (i.e. "This is easy." or "I know how to do this!").

- **Very Positive** – Participant only or primarily had recorded experiences of success in the course during the specified time period (i.e. only earned quality grades, indicative of mastery). Participant was observed to primarily experience success (i.e. showed excitement, made positive comments, positive body language and composure in class). Participant made primarily positive comments about him/herself in relation to

mathematics (i.e. "This is easy." or "I know how to do this!"). Very seldom or no

negative experiences or sentiments were observed.

Method to compare data sources. This section details the methods to assess agreement

between the five data sources for the treatment group: the standard instrument *MSES*, participant

Journals, participant Interviews, participants' course assessments, and the Instructor Perspective.

Each pairing followed a similar method used to assess agreement. The results, descriptions of

agreements and disagreements between the data sources, and a summary will be presented in

Chapter 4. The numbered points below clearly state each pairing:

1. *MSES* and Interviews

2. *MSES* and Journals

3. *MSES* and Course Assessments

4. *MSES* and Instructor Perspective

5. Interviews and Journals

6. Interviews and Course Assessments

7. Interviews and Instructor Perspective

8. Journals and Course Assessments

9. Journals and Instructor Perspective

10. Course Assessments and Instructor Perspective

It is important to note that agreement/alignment between data sources was present if there

was agreement between the two data sources in terms of the participants' upward, downward, or

constant trends in self-efficacy changes, whether indicated by Journals, Interviews, *MSES*

survey, participants' course assessments, or the Instructor Perspective. These trends are what will

be referred to as the story told by each data source because it is a representation of each

participant's experience. Further, "Strongly Agree" indicates an exact or close to exact match between the trends of the two data sources being compared. "Somewhat Agree" indicates a similar overall trend between the two data sources being compared, but with some fluctuations within the data that do not support a strong agreement. "Somewhat Disagree" indicates a weak, opposite overall trend between the two data sources being compared. "Strongly Disagree" indicates an exact opposite overall trend or an inverse relationship between the trends of the two data sources being compared.

For each of the ten pairings, the following three-step method of analysis was conducted to assess agreement. A specific example for the pairing of Interviews and Journals will illustrate this in a clear manner.

- Step 1: To assess agreement, I initially compared the story told by the ratings of the five Journals to the story told by the ratings of the two Interviews. If these two stories about the same participant aligned in terms of the direction of the trend (i.e. gain, no change, or loss) of any perceived changes in self-efficacy, I initially stated "Agree". For example, the Journals from Participant 5 indicate that she began the semester with a Neutral rating, then made gains to Moderately Positive, and spent the second half of the semester with a Very Positive Rating. When compared to the Interviews, this student began the semester with a Neutral rating and ended with a Very Positive rating. Both stories illustrate the same trend of gain in self-efficacy from a Neutral to Very Positive rating and so this is an example of when the two stories agreed. If the two stories did not align, I initially stated "Disagree". All disagreements are described in Chapter 4. This assessment was done 19 times, once for each of the participants for each of the 10 pairings.

- Step 2: After the initial step, I went back through the comparisons to see if I still agreed with the initial stated levels of agreement for each participant. I made note of any changes I made.

- Step 3: I, then, felt it was important to indicate the level or strength of agreement or disagreement because some relationships appeared to be stronger than others. These indicators included "somewhat" and "strongly". These are described earlier in this section.

To help illustrate how the different data sources align in a visual manner, the scales for each are displayed below, in alignment with one another, in Figure 4. Specifically, a score of 0 (No Confidence at all) would correspond to Very Negative on the scales for Journals and Interviews and Instructor Perspective. A score of 1, 2, or 3 (Very little Confidence) would correspond to Moderately Negative and Somewhat Negative on the scale for Journals and Interviews and Mostly Negative, and into Very Negative, on the scale for Instructor Perspective. A score of 4 or 5 (Some Confidence) would correspond to Neutral on the scale for Journals and Interviews and Mixed on the scale for Instructor Perspective. A score of 6 or 7 (Much Confidence) would correspond to Somewhat Positive and into Moderately Positive on the scale for Journals and Interviews and Mostly Positive on the scale for Instructor Perspective. A score of 8 or 9 (Complete Confidence) would correspond to Moderately Positive into Very Positive on the scale for Journals and Interviews and Very Positive on the scale for Instructor Perspective. Being able to use and create ordinal scales for qualitative data, as advised from close work with my advisors, enabled this study to have a relatively principled way to compare and discuss the results of data analysis for multiple sources of data. However, this causes the descriptive richness of the data to lessen, especially for students who had only agreements between the results of their

data sources. This richness might have been preserved from choosing a different direction for data analysis, such as individual case studies. The participants for whom any disagreements resulted will be discussed in more detail in response to Research Question 4, and so the richness of data is better maintained for these participants.

Figure 4: Comparison of Data Source Scales
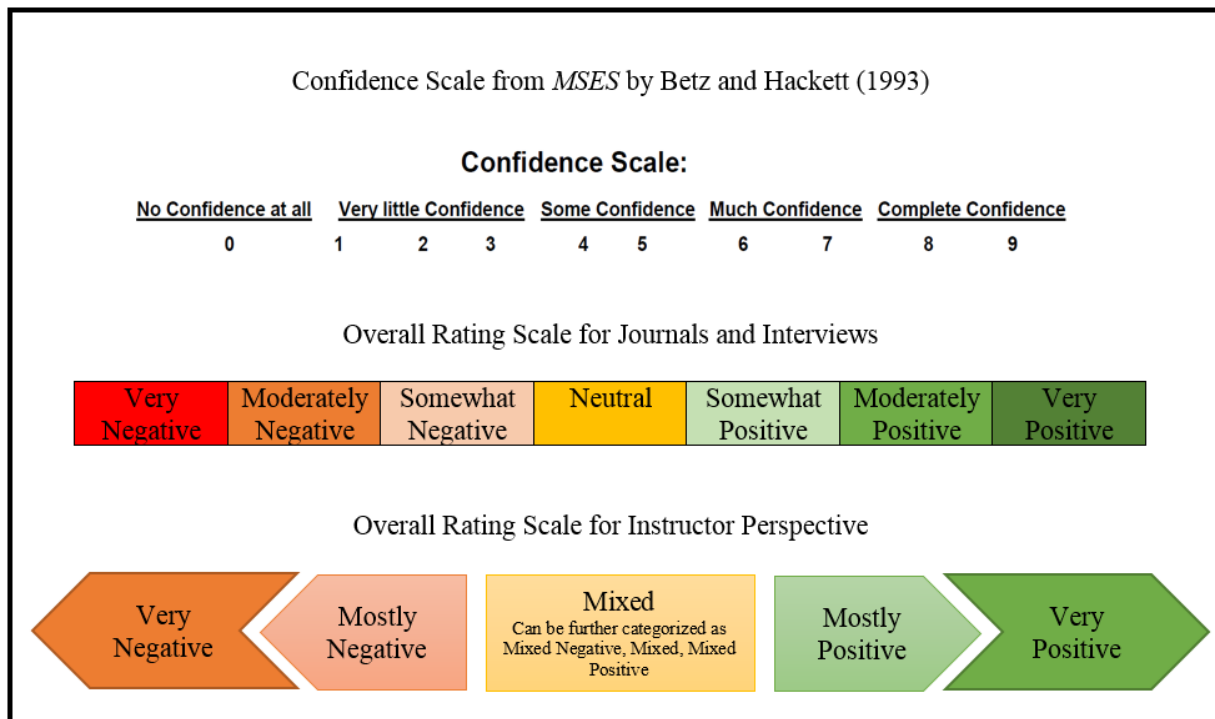
Following the initial set of ten comparisons between data sources, three additional rounds of comparison occurred. The reason for this was to check and lessen the bias that I have as an instructor and researcher. The second round resulted in only seven adjustments, but all within the agreement categories (i.e. adjusting from Strongly Agree to Somewhat Agree or vice versa). The third round resulted in no adjustments. As a final check, the fourth round resulted in twenty-three adjustments out of a total of 190 comparisons between the data sources, but, again, all were within agreement categories except for one change from Agree to Disagree. Checking these

results multiple times was important because I wanted to check and lessen any bias and ensure that the results of the comparisons was as accurate as possible.

Once the agreements and disagreements between data sources were identified, it was then important to see if any outlier data sources existed. To determine this, I created the following diagram in Figure 5 for each participant. This diagram is one that would result for a participant that had agreement exist between all data sources. Each line represents agreement. The absence of a line between data sources means that disagreement exists. Representing the agreements and disagreements in a visual way helped to identify any data sources that could be considered an outlier for an individual participant. These will be presented in Chapter 4.

Figure 5: Example of the Relationships of Data Sources for Participants Whose Results Agree
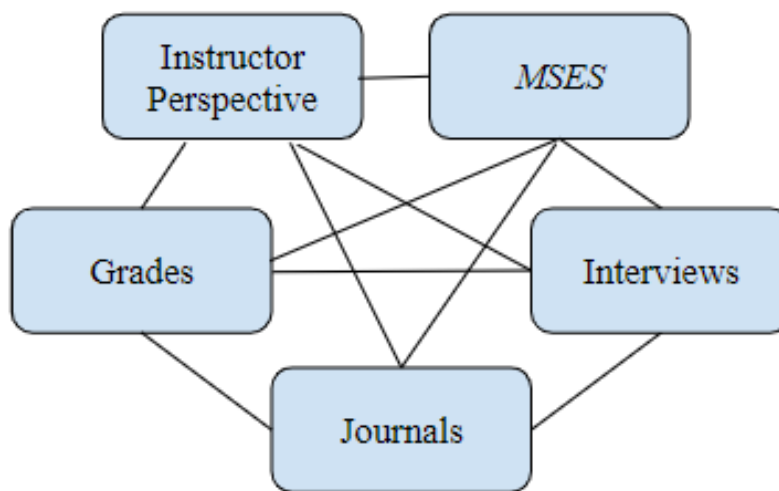
Once these diagrams have been presented, the bulk of the discussion for this research question will focus on participants who have disagreements between their data sources. For each of these participants, a detailed description of the student and his/her perceived self-efficacy, with evidence from their data sources, will be provided. In addition, these participants will be considered together, along with each student's unique circumstances, to inform whether patterns exist when making sense of the occurrence of disagreements for some students and not others.

Inter-Rater Reliability

It is important to establish inter-rater reliability for the current study because I want to ensure that the codes I assign to the collected data are clear and would likewise be assigned by another person. A colleague of mine at Michigan State University agreed to assist me in this step of my work. We met virtually for an initial training session after I sent him my codebook used for coding Journals and Interviews (Table 3), a de-identified data set for Journal 2, and a blank results table for him to record his codes and notes. I chose Journal 2 because this was administered during Week 5 during Fall 2019 and it is likely that the participants were more comfortable at this time rather than during Journal 1. Additionally, the questions in Journal 2 were more similar to those asked in later Journals. Journal 4 was also used for inter-rater reliability coding because it is later in the semester than Journal 2 and had an additional component of reflecting on a past exam score.

The goal was to have an inter-rater reliability of at least 80% upon the conclusion of this phase because Creswell (2014) states that "the consistency of the coding be in agreement at least 80% of the time for good qualitative reliability" (p. 203). We completed three rounds of coding for inter-rater reliability, the results of which are depicted in Table 6. I would like to direct the reader's attention to the row titled *Agreed upon codes out of commonly coded statements (a subset of the row above it)* because these percentages show the number of statements that we both chose to code and our level of agreement on those. I highlight this row because Creswell (2014) states, "It is not that [the coders] code the same passage of text but whether another coder would code it with the same or a similar code" (p. 203). These percentages show that we agreed on a statement's code over 90% of the time for each round of coding and agreed 100% of the time after discussion. Further, it is also important to discuss how our agreements contributed to

78

the variation in the percentages for *Ratings*. These are further detailed in the row below showing that our ratings were not far off one another. Two contributing issues to these lower percentages were (a) how we each chose to split a statement, which will be further discussed, and (b) that some statements were coded but not chosen by both coders. Both issues, along with our high level of agreement, contribute to the observed percentages for *Ratings* below.

Table 6: Summary of Rounds to Establish Inter-rater Reliability

		Round 1 1st half Journal 2	Round 2 2nd half of Journal 2	Round 3 All of Journal 4
Pre-Discussion	**Ratings for each journal for an individual participant**	7/10 70%	6/9 66% *can improve based on splitting statements	9/19 47% *can improve based on splitting statements
		On the other 3/10, we were only off by 1 rating.	On the other 3/9, we were only off by 1 rating.	On the other 10/19, 5 were off by 1 rating. 4 were off by 2 ratings. 1 was off by 3 ratings.
	Agreed upon codes out of total coded statements	40/55 72.7%	57/78 73%	106/142 75%
	Agreed upon codes out of commonly coded statements (a subset of the row above)	40/44 91%	57/60 95%	106/117 90.5%
Post-Discussion	**Ratings**	100%	100%	100%
	Total Statements	100%	73/78 = 93.6% *we agreed to consult with advisors on the "help" and "questions" statements	100%
	Chosen by both	100%	100%	100%

Introduction and training. In our initial meeting, I explained the codebook with examples, the coding process, and how to assign a rating to each Journal. My colleague coded the first page in the Journal 2 data set (Participant 16's Journal 2 submission) as a sample piece on his own during our initial meeting. We discussed questions he had with the coding process and came to an agreement on all codes from this sample piece.

Round 1. Following our initial meeting, my colleague coded the first half of participant

Journals (9 Journals) in the Journal 2 data set (of the total 19 Journals). I compared the ratings

first and then I went line by line to compare every coded statement that each of us identified. The

results of this is shown in Table 6 in the *Pre-Discussion* rows for the *Round 1* column. We met

again to discuss statements that we disagreed on (four statements) or statements that one of us

chose to code and the other did not (11 statements). We came to an agreement on all 15

statements, which is reflected in the *Post-Discussion* rows for the *Round 1* column.

After our meeting, I counted the coded statement results for each participant and

determined if a new rating should be assigned. This resulted in a change of two ratings from my

original codes and one change in my colleague's original codes. I also determined that a couple

changes in my own coding scheme must be made in order to reflect the new perspective that I

gained after this meeting. I applied these changes to all Journals and Interviews. Details of these

changes are noted in Appendix J.

After applying this new perspective to the data, I made four changes to statements in my

coding to the second half of Journal 2, but those changes had no effect on the ratings. Continuing

through the rest of the data, 13 additional changes to individual statements' codes were made.

These 13 changes did affect the results of five ratings in the other Journal sets.

In this round of coding my bias as both the instructor and researcher became apparent.

For example, I observed that one participant appeared to have a big ego while in my classroom,

which led me to code that participant's statement, "I felt like it was too easy at a certain point.

But then I got challenged." (Journal 2, p. 5) as Variable because the student indicated a variation

in his experience about the difficulty level of the course. My colleague did not know this about

this participant when coding, and so he coded the statement as Positive – Present. We discussed

this statement in detail for some time and agreed on coding the statement as Variable. I feel that this bias can be helpful in bringing out the true meaning of the statement that students write, especially with written data. This bias, however, has also made it difficult to separate myself from the data and participants. No changes to my coding perspective were made.

Table 7 shows both of our initial ratings for Round 1 and is listed here to show how they relate to the final rating. Our discussion and decisions about statements together is what determined the final ratings. Similar results tables for Rounds 2 and 3 can be found in Appendix J.

Table 7: Round 1 for First Half of Journal 2

Participant	My Rating	Colleague's Rating	Agreed Upon Final Count and Rating
Participant 4	Somewhat Positive (2-4-1)	Moderately Positive (2-6-1)	**Moderately Positive (2-6-1)**
Participant 16	Neutral (2-3-1)	Neutral (2-3-1)	**Neutral (2-3-1)**
Participant 2	Very Positive (0-2-1)	Very Positive (0-4-0)	**Very Positive (0-3-1)**
Participant 18	Somewhat Positive (2-4-0)	Somewhat Positive (2-4-0)	**Somewhat Positive (2-4-0)**
Participant 9	Somewhat Positive (2-4-0)	Neutral (2-3-0)	**Somewhat Positive (2-4-0)**
Participant 1	Neutral (1-1-1)	Neutral (1-1-2)	**Neutral (1-1-2)**
Participant 17	Somewhat Positive (1-2-0)	Somewhat Positive (1-2-0)	**Somewhat Positive (1-2-0)**
Participant 10	Moderately Positive (1-3-0)	Very Positive (0-4-1)	**Very Positive (0-4-1)**
Participant 19	Neutral (2-3-0)	Neutral (2-3-0)	**Neutral (2-3-0)**
Participant 15	Moderately Negative (3-1-1)	Moderately Negative (3-1-1)	**Moderately Negative (3-1-1)**

Round 2. Following the same process from Round 1, we met to discuss statements that we disagreed on (three statements) or statements that one of us chose to code and the other did not (18 statements). The results of this is shown in Table 6 in the *Pre-Discussion* rows for the *Round 2* column. In Round 2, there were instances of both of us coding a statement with the same overall code (i.e. Positive - Present) but one of us would break the statement into two parts and give it two positive codes while the other would not break the statement up and assign one positive code. There were fairly equal instances of this happening between the two of us and it happened 10 times. Even though we agreed on the code for the statement, it affected our ratings for a participant's Journal by one level in two cases. Since we agreed on the code that was assigned, we did not feel there was an issue to be resolved. After consulting with my advisors, it was deemed appropriate to include a table that illustrates the overall code assigned to each affected Journal and how the differences in parsing could change those ratings. For reference, see Appendix J and in particular Table 36, which shows that of the five Journals that were affected by splitting in this round, the ratings of three Journals remained the same and two Journals showed a change by one level. Therefore, the splitting of statements does not appear to have a significant effect on final ratings.

In addition, discussions about coding that altered the codebook for future coding sessions included (a) participants' use of the word *help* in their journals and (b) when participants used the phrase "I have questions about…" a mathematical topic. The first issue was coded as positive as long as the participant has done one of the following (a) specifically connected a reason for how something/someone was helpful to themselves in math or (b) has made an explicit connection to him/herself. If they did not do this, then the statement was left un-coded. The second issue was left open for interpretation. If the participant indicated that what they have

questions about is causing them to struggle, then it was coded as negative. However, if it is a simple statement that s/he has questions about a topic with no other information, it should not be coded. More detail about these determinations can be found in Appendix J.

Round 2 discussions resulted in a change of two ratings from my original codes and two changes for my colleague's codes. Following the conclusion of Round 2 of establishing inter-rater reliability, I determined that the perspective with which I coded the Journals and Interviews at this point in time was accurately represented. I made this decision because the conversations I had with my advisors and colleague supported my thinking and the perspective that had already been applied.

Round 3. Even though it seems that inter-rater reliability was established after Round 2, I felt that one additional round would be beneficial to the study and solidify our strong agreement on statements that we both chose to code (Creswell, 2014). Following the same process as Rounds 1 and 2, we met again to discuss statements that we disagreed on (11 statements) or statements that one of us chose to code and the other did not (25 statements) from the coding results of Journal 4. The results of this is shown in Table 6 in the *Pre-Discussion* rows for the *Round 3* column. In this round, the issue of the division of statements arose again (recall, one of us would code a statement once and the other would split it in two parts and assign the same code twice). The only new issue that arose, which was one in an extremely specific context, was in relation to participants' expectations versus reality of an exam score that they received back. For example, participants would express self-doubt followed by excitement because they did not think they were going to perform as well as they did on an exam (and vice versa). My colleague would code this as Variable, and I would, code the self-doubt portion as negative and

the excitement portion as positive. Ultimately, we decided to code these as Variable because it best captures the change a participant underwent.

After our meeting, I counted the coded statement results for each participant and determined if a new rating must be assigned. This resulted in a change of five ratings from my original ratings and eight changes for my colleague. Upon conclusion of this round, I decided that it was worthwhile to revise my codes in the other Journals for statements about expectations and reality with their latest exam score. The Interviews were not affected by statements of this type. The results of this only affected Journals 3 and 5 because Journals 1 and 2 did not ask students to reflect on an exam score and Journal 4 was already addressed. The revisions in coding to reflect this new, and final, perspective for the current study made 21 changes to the codes of individual statements across the two Journals. These affected the results of 12 ratings in Journals 3 and 5.

The next chapter depicts the results of the previously described data collection and data analysis methods for the participants in the control and treatment groups. The results are presented in order of the research questions that drives the current study. A discussion of these results will follow.

CHAPTER 4: RESULTS

Research Question 1: Changes in Self-Efficacy Measured with Standard Instrument

Research Question 1 is: "What, if any, changes in mathematics self-efficacy occur for at-risk college students, as measured with a standard instrument, when the instructional treatment is implemented?" Upon completion of the semester-long mathematics course, each of the nineteen students illustrated some gain, whether great or small, in their self-efficacy as portrayed by their scores on the standard instrument (*MSES*). The recruitment of these participants is discussed in Chapter 3. Recall the scale by Betz and Hackett (1993), which is used to interpret the results of the instrument and is displayed in Figure 6 below.

Figure 6: Confidence Scale from *MSES* by Betz and Hackett (1993)

Confidence Scale:

No Confidence at all	Very little Confidence	Some Confidence	Much Confidence	Complete Confidence
0	1 2 3	4 5	6 7	8 9

Total score for treatment group. Table 8 represents the scores of the 19 participants at three time points throughout the semester (beginning of the semester: *MSES* 1, middle of the semester: *MSES* 2, and end of the semester: *MSES* 3). These scores are out of 9 points and are calculated from the average of each participant's responses to 34 items on the survey. For each item, the participant chose a response from 0 (No Confidence at all) to 9 (Complete Confidence). The choice to use three points of time instead of two was to observe whether there was a greater change in the first half of the semester or the second half of the semester, in addition to considering overall change.

Table 8: *MSES* Scores for Treatment Group

Participant	*MSES* 1	*MSES* 2	*MSES* 3
1	4.15	4.65	7.59
2	7.82	8.15	8.36

Table 8 (cont'd)

3	5.53	5.61	6.25
4	5.32	6.82	8.03
5	4.97	5.53	5.65
6	4.15	5.50	6.03
7	4.38	4.85	8.12
8	2.79	5.12	5.35
9	4.18	5.44	6.12
10	4.97	4.53	7.53
11	4.85	6.32	6.50
12	3.64	5.18	6.26
13	2.29	4.26	6.91
14	3.71	3.82	4.30
15	6.09	6.53	7.71
16	5.85	7.71	7.44
17	6.50	6.82	7.88
18	4.26	6.79	6.94
19	5.59	6.24	7.41
Mean	**4.79**	**5.78**	**6.86**

The class mean at the beginning is 4.79 points (Some Confidence), at the middle is 5.78 points (Some Confidence), and at the end is 6.86 points (Much Confidence). In addition, the standard deviation at the beginning is 1.31, at the middle is 1.17, and at the end is 1.08. The range of scores at the beginning is 2.29 – 7.82 (a difference of 5.53 points) , at the middle is 3.82 – 8.15 (a difference of 4.33 points), and at the end is 4.30 – 8.36 (a difference of 4.06 points). Notice that the standard deviations and ranges become smaller as the semester progresses. The smaller standard deviation and range over time indicate that the scores, overall, are getting closer to the mean with less variation. This is important to mention because it illustrates the changes that the participants as a group experience throughout the semester, and that as time progresses there is less variation in the scores on this instrument. The average point gain, or change, for the entire group over the course of the semester from *MSES* 1 to *MSES* 3 is 2.07 points. The range of gains in scores across the semester suggests that the treatment was effective in increasing participants' self-efficacy, no matter the level of preliminary score on *MSES* 1. This, however, will be further scrutinized with quantitative data analysis. Further, the largest overall gain (*MSES* 1 to *MSES* 3) is 4.62 points for Participant 13 and the smallest overall gain is 0.54 points for

Participant 2. Two students (Participants 10 and 16) show a loss at one time, but still illustrate an overall gain.

A visual representation of the scores for the treatment participants is displayed in Figure 7. The shift in self-efficacy that the participants make as a group can be easily observed. With each administration, it is clear that the average height of the bars increases, which indicates that self-efficacy (as measured by the *MSES*) increases over time. In addition to looking at each set of bars for each administration as a whole, responses from individual participants can be tracked.

Figure 7: Bar Graph of Treatment Participant Scores for *MSES*
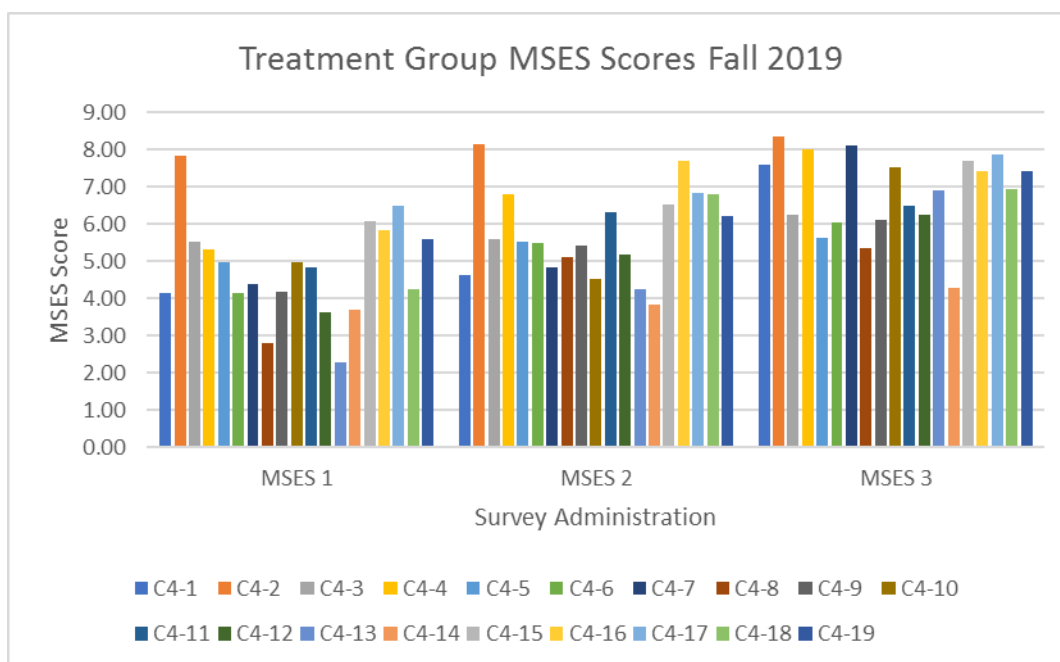
I conducted a Repeated Measures ANOVA in SPSS with the results of the treatment group to test for differences in mean scores across time. Even though this data is ordinal in nature, this test has shown to produce robust results in the face of potential violations of assumptions posed by the use of Likert scales, even with small sample sizes (Norman, 2010). I used time as the independent variable and the three *MSES* results were used as the dependent variable. Because I have three time points for the treatment group, I used dummy codes to

compare *MSES* 1 to *MSES* 2 to see if there was an improvement in the first half of the semester, and to compare *MSES* 2 to *MSES* 3 to see if there was an improvement in the second half of the semester. Three hypotheses were tested using Repeated Measures ANOVA (twice for the treatment group and once for the control group). The control group only had two time points, and so a separate Repeated Measures ANOVA compared their scores on *MSES* 1 to *MSES* 3. I did a Bonferroni correction and used an alpha level of .05/3 = .017 when considering and reporting my results. Table 9 shows the results of the Repeated Measures ANOVAs. I will discuss the results of the treatment group now, but I will discuss the comparison of the two groups in the next section for Research Question 2.

Table 9: Results of Repeated Measures ANOVAs Testing Changes in Self-Efficacy Over Time

	F	*df*	*p*	Partial Eta Squared
Treatment Group (*MSES* 1-*MSES* 2)	26.81	1,18	< .001	.60
Treatment Group (*MSES* 2-*MSES* 3)	18.66	1,18	< .001	.51
Control Group (*MSES* 1-*MSES* 3)	21.19	1,42	< .001	.34

There was a significant effect of time on *MSES* scores, Wilks' Lambda = 0.20, $p < .001$. As shown in Table 9, the dummy codes indicate significant differences between *MSES* 1 to *MSES* 2 and *MSES* 2 to *MSES* 3. Recall the treatment group's mean *MSES* scores at each time point: $M_{MSES1} = 4.79$, $M_{MSES2} = 5.78$, and $M_{MSES3} = 6.86$. There was a statistically significant effect of time on *MSES* scores for the treatment participants in the first half of the semester, $F(1, 18) = 26.81$, $p < .001$ and also in the second half of the semester, $F(1, 18) = 18.66$, $p < .001$. The results suggest that participants in the treatment group exhibited significant increases in self-efficacy, with continued growth across both halves of the semester.

Subscale scores for treatment group. To narrow the focus of the results of the *MSES*

scores as a total score, Betz and Hackett (1993) also encouraged consideration of the results in

terms of two subscales. The first subscale (Part 1: Everyday Math Tasks) consists of 18 items:

"How much confidence do you have that you could successfully:" followed by 18 varied math

tasks (e.g. estimating their grocery bill in their head as they shop). The second subscale (Part 2:

Math Courses) consists of 16 items: "Please rate the following college courses according to how

much <u>confidence</u> you have that you could complete the course with a <u>final grade</u> of 'A' or 'B'"

(e.g. Accounting). Table 10 shows the results of the subscales for the treatment group across the

three time periods.

Table 10: *MSES* Scores for Two Subscales for Treatment Group

Participant	Part 1: Everyday Math Tasks			Part 2: Math Courses		
	MSES 1	*MSES* 2	*MSES* 3	*MSES* 1	*MSES* 2	*MSES* 3
1	4.00	4.72	7.78	4.33	4.56	7.38
2	7.83	8.56	8.44	7.81	7.69	8.27
3	5.78	6.22	6.78	5.21	4.87	5.57
4	5.22	8.00	8.67	5.44	5.40	7.31
5	5.83	6.28	6.89	3.93	4.69	4.25
6	4.78	6.89	7.11	3.44	3.94	4.81
7	4.78	5.44	8.56	3.94	4.19	7.63
8	3.28	5.17	4.89	2.25	5.06	5.88
9	4.50	6.56	6.83	3.81	4.19	5.31
10	4.72	6.50	7.11	5.25	2.31	8.00
11	5.39	7.44	8.33	4.25	5.06	4.44
12	3.50	5.33	6.39	3.80	5.00	6.13
13	2.67	6.06	7.89	1.88	2.25	5.81
14	4.44	5.67	6.33	2.88	1.75	1.87
15	5.61	6.89	7.78	6.63	6.13	7.63
16	6.17	8.06	8.56	5.50	7.31	6.19
17	6.11	7.56	8.44	6.94	6.00	7.25
18	3.72	6.61	7.00	4.88	7.00	6.88
19	5.28	6.67	7.94	5.94	5.75	6.81
Mean	4.93	6.56	7.46	4.64	4.90	6.18

Increases in the group's mean scores on the two subscales are evident, but noticeably

greater gains are made for Part 1 than Part 2. Considering Part 1 (Everyday Math Tasks), all

participants illustrated overall gains in their self-efficacy. The class mean at the beginning is 4.93

points (Some Confidence), at the middle is 6.56 points (Much Confidence), and at the end is 7.46

points (Much Confidence). In addition, the standard deviation for Part 1 at the beginning is 1.21,

at the middle is 1.04, and at the end is 1.00. The range of scores for Part 1 at the beginning is

2.67 – 7.83 (a difference of 5.16 points) , at the middle is 4.72 – 8.56 (a difference of 3.84

points), and at the end is 4.89 – 8.67 (a difference of 3.78 points). Overall, the average point gain

is 2.53 points, with greater gains made during the first half of the semester than the second half

of the semester. When the Part 1 subscale scores are compared to the total *MSES* scores, the

average scores for the group are higher for Part 1 at every administration. Considering individual

participants, the largest overall gain (*MSES* 1 to *MSES* 3) is 5.22 points for Participant 13 and the

smallest overall gain is 0.61 points for Participant 2. These are the same two participants who

had the highest and lowest gains in *MSES* total scores without considering subscales. Two

students (Participants 2 and 8) show a loss between the second and third administrations, but still

retain an overall gain.

Shifting our focus to Part 2 (Math Courses), reveals that more variation in scores across

participants are present. There is an average gain of 1.54 points across the whole group, with all

but one participant showing an overall gain. The class mean at the beginning is 4.64 points

(Some Confidence), at the middle is 4.90 points (Some Confidence), and at the end is 6.18 points

(Much Confidence). In addition, the standard deviation for Part 2 at the beginning is 1.56, at the

middle is 1.62, and at the end is 1.58. The range of scores for Part 2 at the beginning is 1.88 –

7.81 (a difference of 5.93 points) , at the middle is 1.75 – 7.69 (a difference of 5.94 points), and

at the end is 1.87 – 8.27 (a difference of 6.40 points). Greater gains are made during the second

half of the semester compared to the first half of the semester. When compared to the total *MSES*

scores and Part 1 scores, the average scores for the group are lower for Part 2 at every

administration. All participants illustrated an overall gain except Participant 14, who exhibited a decline in self-efficacy for Math Courses from beginning to end. The largest overall gain (*MSES* 1 to *MSES* 3) is 3.94 points for Participant 13, and the only loss for the group is -1.01 points for Participant 14. The *MSES* indicates only increasing scores for the following participants: 1, 6, 7, 8, 9, 12, 13. The remaining eleven participants experience a combination of increase and decrease in their self-efficacy toward Math Courses, but still report overall gains from beginning to end.

In summary, all participants in the treatment group made overall gains in their total self-efficacy scores on the *MSES* when comparing the first administration of the survey to the last administration of the survey upon completion of the treatment course. On average, participants in the treatment group gained 2.07 points on their overall score, 2.53 points on their Part 1 subscale score, and 1.54 points on their Part 2 subscale from *MSES* 1 to *MSES* 3. The average overall scores on the three administrations indicated fairly equal gains made during both halves of the semester. However, examining the two subscales separately revealed that students made greater gains in confidence for Everyday Math Tasks in the first half of the semester and greater gains in confidence for Math Courses in the second half of the semester. It is important to note, however, that two participants did experience a decline in their total score at one point. One participant between the first and second administrations and a second participant between the second and third administrations. Both participants, in addition to all other participants in the treatment group, still showed overall gain from start to finish.

Research Question 2: Comparisons of the Control and Treatment Group

Research Question 2 is: "How do these results compare to students of similar placement, but who do not receive the treatment, as measured with the same instrument?" To further

understand the results of the treatment group, a comparison to a larger control group will now be considered. The recruitment of these participants is discussed in Chapter 3. This control group, who did not receive any treatment, took the survey twice. This is the same survey that was taken by the treatment group. The first administration of the survey (*MSES* 1) was at the beginning of the semester. The second administration of the survey (*MSES* 3) was at the end of the semester, after their standard non-treatment algebra course was completed. Table 11 lists the scores of 53 participants on the *MSES* survey (only 43 completed the second survey).

Table 11: *MSES* Scores for Control Group

Participant	*MSES* 1	*MSES* 3		Participant	*MSES* 1	*MSES* 3
nonC4-1	4.47	-		nonC4-28	2.79	-
nonC4-2	2.91	4.24		nonC4-29	5.44	4.68
nonC4-3	7.71	8.35		nonC4-30	5.38	-
nonC4-4	7.88	-		nonC4-31	6.88	7.44
nonC4-5	6.91	7.82		nonC4-32	6.03	7.59
nonC4-6	6.41	7.00		nonC4-33	3.94	4.79
nonC4-7	4.84	-		nonC4-34	5.97	4.41
nonC4-8	4.39	7.38		nonC4-35	3.24	4.00
nonC4-9	4.26	5.91		nonC4-36	3.09	6.21
nonC4-10	4.15	5.15		nonC4-37	5.50	5.88
nonC4-11	5.74	7.09		nonC4-38	5.00	-
nonC4-12	4.26	5.29		nonC4-39	6.29	7.47
nonC4-13	3.97	6.73		nonC4-40	2.41	-
nonC4-14	4.97	4.71		nonC4-41	4.35	5.29
nonC4-15	5.12	4.94		nonC4-42	7.00	7.41
nonC4-16	3.26	5.35		nonC4-43	4.76	4.03
nonC4-17	5.56	5.76		nonC4-44	2.41	5.62
nonC4-18	2.68	3.32		nonC4-45	5.74	-
nonC4-19	6.38	5.82		nonC4-46	2.94	5.24
nonC4-20	3.06	-		nonC4-47	5.47	4.26
nonC4-21	2.47	6.09		nonC4-48	5.82	5.24
nonC4-22	3.41	6.00		nonC4-49	6.15	5.68
nonC4-23	6.65	-		nonC4-50	3.71	4.06
nonC4-24	3.91	5.29		nonC4-51	5.03	5.62
nonC4-25	3.24	4.44		nonC4-52	5.47	5.65
nonC4-26	4.56	4.91		nonC4-53	4.74	4.53
nonC4-27	6.35	6.85		**Mean**	4.81	5.66

Total score for control group. As can be seen in Table 11, the *MSES* scores indicate a gain in self-efficacy in mathematics for 33 participants and a loss for 10 participants. The mean

score for the first administration for the control group is 4.81 points (Some Confidence). The mean score for the last administration is 5.66 points (Some Confidence), which is a difference of 0.85 points between the two administrations. The standard deviation for the first administration is 1.44 and for the last administration is 1.21. The range of scores at the first administration is 2.41 – 7.88 (a difference of 5.47 points) and at the last administration is 3.32 – 8.35 (a difference of 5.03 points).

Please note that the remaining 10 participants either dropped the course ($n = 2$), withdrew from the university ($n = 5$), or did not complete a second survey ($n = 3$). These 10 participants have dashes in place of a *MSES* 3 score in Table 11. Control participants who did and did not complete the study were compared using a Chi-Square Test of Independence on (a) gender, χ^2 (1, $N = 53$) = .09, $p = .769$, (b) race, χ^2 (2, $N = 53$) = 2.55, $p = .280$, and (c) urban/rural high school, χ^2 (1, $N = 53$) = .59, $p = .442$; no differences were significant. They were also compared using an Independent Samples Test (T-test) for (a) high school GPA, $t(51) = 1.98$, $p = .053$, (b) Total SAT scores, $t(51) = -.79$, $p = .431$, (c) Math Section SAT scores, $t(51) = -1.80$, $p = .077$, and (d) *MSES* 1 scores, $t(51) = -.02$, $p = .982$; no differences were significant.

I conducted a Repeated Measures ANOVA in SPSS with the results of the control group to test for differences in mean scores across time with a Bonferroni correction. Recall that Table *9* presented the results of this test. The 43 participants who completed both administrations of the *MSES* were considered for the results of the control group. I used time as the independent variable and the two *MSES* results were used as the dependent variable. The results indicate there was a statistically significant effect of time on *MSES* scores for the control participants across the semester, Wilks' Lambda = .67, $F(1, 42) = 21.19$, $p < .001$. These results suggest that participants in the control group exhibited significant increases in self-efficacy across the whole

semester.

Total score comparison of both groups. I will now illustrate that the results of the mean scores for both groups have significant differences in favor of the treatment group, in addition to showing how comparable the two groups were prior to the treatment being administered. For reference, Table 12 shows the treatment and control group mean scores and standard deviations on the *MSES* at the beginning and end of the semester. Figure 8 shows the treatment and control group means scores on the *MSES* at the beginning and end of the semester. Additionally, Table 1 contains both groups' demographic information including High School GPA, Math Section SAT score, and SAT total score, which is the information used for the Independent Samples Test described below.

Table 12: Comparison of Treatment and Control *MSES* Scores

	MSES 1		*MSES* 3	
	Mean	St. Dev.	Mean	St. Dev.
Treatment Group	4.79	1.31	6.86	1.08
Control Group	4.81	1.44	5.66	1.21

Figure 8: Comparison of Treatment and Control Group Total *MSES* Score
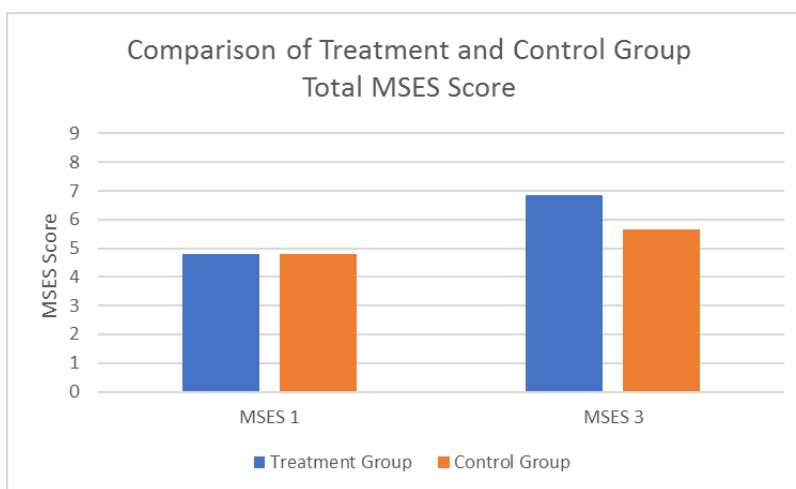
I conducted an Independent Samples Test (T-test) in SPSS in order to compare the

treatment and control groups on the bases of High School GPA, Math Section SAT, SAT Total,

MSES 1, *MSES* 1 Part 1 subscale, and *MSES* 1 Part 2 subscale at the beginning of the semester.

Since six hypotheses were tested using this T-test, I did a Bonferroni correction and used an

alpha level of .05/6 = .008 when considering and reporting my results.

Table 13 shows the results of the Independent Samples Test (T-test). There were no

significant differences between the two groups at the start of the semester when considering High

School GPA ($t(70) = -.04$, $p = .965$), Math Section SAT ($t(70) = -.07$, $p = .942$), the first *MSES*

administration ($t(70) = .06$, $p = .955$), Part 1 subscale ($t(70) = 1.23$, $p = .223$), and Part 2 subscale

($t(70) = -1.07$, $p = .290$). This shows that the two groups are comparable based on these

measures. The control group, however, has SAT Total scores that are on average higher than

those of the treatment group, $t(70) = 3.31$, $p = .001$ ($M_{Treatment} = 861.05$, $M_{Control} = 924.62$) and

will be discussed as a limitation to this study. In addition, and as indicated in Chapter 3, the

treatment and control groups were compared using a Chi-Square Test of Independence on (a)

gender, χ^2 (1, $N = 72$) = 1.07, $p = .300$, (b) race, χ^2 (2, $N = 72$) = 1.51, $p = .470$, and (c)

urban/rural high school, χ^2 (1, $N = 72$) = 1.73, $p = .188$; no differences were significant.

Even though no differences were detected from statistical inference, it is important to

highlight any observable differences between the participants who made up the two groups. First,

when considering gender, 52.6% of the treatment group participants were female compared to

66.0% in the control group. Second, when considering race, 15.8% were Black, 5.3% were

Hispanic, and 78.9% were White from the treatment group, while 30.2% were Black, 3.8% were

Hispanic, and 66.0% were White from the control group. Third, when considering participants'

high school demographics, 73.7% from the treatment group were from a rural high school

compared to 50.9% from the control group. It seems that the treatment group had noticeably

higher participation from students who were male, White, and/or from rural high schools than

those who participated in the control group, while students who were female, Black, and/or from

urban high schools had higher rates of participation in the control group. It is important to

recognize that differences in the population of the two groups could have contributed to the

results of this study, which will be discussed in Chapter 5.

Table 13: Results of Independent Samples Test to Compare Treatment and Control Groups

	Independent Samples Test			Means	
	t	df	Sig. (2-tailed)	Mean for Treatment	Mean for Control
HS GPA	-.04	70	.965	2.992	2.987
Math Section SAT	-.07	70	.942	444.74	444.06
SAT Total	3.31	70	.001	861.06	924.62
MSES 1	.06	70	.955	4.79	4.81
MSES 1 Part 1	1.23	70	.223	4.93	5.41
MSES 1 Part 2	-1.07	70	.290	4.64	4.15

To consider the end of semester results of the *MSES* 3 scores between the treatment and

control groups, I conducted a Repeated Measures ANOVA in SPSS to test for differences in the

groups' mean scores across time. This test considers the change in scores made by both groups

(from *MSES* 1 to *MSES* 3) and whether those changes are significant, ultimately considering a

time-by-treatment interaction. Since three hypotheses were tested using this test (once for total

MSES and twice for *MSES* subscales), I did a Bonferroni correction and used an alpha level of

$.05/3 = .017$ when considering and reporting my results. This test is important because

previously discussed results have shown that both groups improved across the semester, but it is

unclear whether the size of the improvement is similar. The test of the interaction was

statistically significant, indicating that the treatment participants made greater improvements in

MSES scores across time than those of the control group, Wilks' Lambda = 0.81, $F(1, 60) =$

13.90, $p < .001$. A sensible interpretation of the statistically significant interaction result was that both groups improved over time, but the treatment group's scores on the *MSES* improved more even though both groups had similar starting points on *MSES* 1 ($M_{Treatment} = 4.79$ and $M_{Control} = 4.81$). Therefore, the results of this measure (the *MSES*) imply that the levels of self-efficacy between the two groups were different at the end of the semester; the treatment group ($M_{Treatment} = 6.86$) ended with higher levels of self-efficacy on *MSES* 3 than the control group ($M_{Control} = 5.66$).

Additionally, it is important to discuss the effect size of this study. An online effect size calculator was used because SPSS does not calculate effect size. Further, the result is theoretical for any study. The effect size calculation considers the means and standard deviations of *MSES* 3 for both groups of participants, in addition to the number of participants in both groups. The effect size (Cohen's *d* value) for the difference in *MSES* scores between the treatment and control groups at the end of the semester was 0.96. This value indicates that *MSES* scores in the treatment group were almost an entire standard deviation higher than the control group at the end of the semester, which is considered a large effect size. It is important to reiterate that both groups' self-efficacy improved throughout the semester, but the control group did not improve as much as the treatment group. It appears that something in the treatment had more of an impact on participants' self-efficacy, but it is also possible that factors outside of the treatment influenced these results, which will be discussed in Chapter 5.

Subscale scores comparison of both groups. We can also compare the scores of the two subscales for both groups. The *MSES* subscale scores on Part 1 and Part 2 for the control group are given in Table 14 and those for the treatment group are provided again in Table 15.

Table 14: *MSES* Scores for Two Subscales for Control Group

Participant	Part 1		Part 2		Participant	Part 1		Part 2	
	MSES 1	*MSES 3*	*MSES 1*	*MSES 3*		*MSES 1*	*MSES 3*	*MSES 1*	*MSES 3*
nonC4-1	5.22	-	3.63	-	nonC4-28	3.28	-	2.25	-
nonC4-2	3.33	6.06	2.44	2.07	nonC4-29	6.78	6.78	3.94	2.31
nonC4-3	8.17	8.72	7.19	7.94	nonC4-30	4.78	-	6.14	-
nonC4-4	7.67	-	8.13	-	nonC4-31	7.44	7.67	6.25	7.19
nonC4-5	7.00	8.17	6.81	7.44	nonC4-32	7.00	8.94	4.94	6.06
nonC4-6	6.50	7.06	6.31	6.94	nonC4-33	5.28	6.33	2.44	3.06
nonC4-7	4.00	-	5.93	-	nonC4-34	7.61	6.06	4.13	2.56
nonC4-8	2.83	7.22	6.27	7.56	nonC4-35	3.61	4.28	2.81	3.69
nonC4-9	4.89	6.33	3.56	5.44	nonC4-36	2.44	5.61	3.81	6.88
nonC4-10	3.50	5.83	4.88	4.38	nonC4-37	6.67	7.39	4.19	4.19
nonC4-11	6.06	7.39	5.38	6.75	nonC4-38	6.56	-	3.25	-
nonC4-12	5.11	6.39	3.31	4.06	nonC4-39	8.00	8.00	4.38	6.88
nonC4-13	5.06	8.22	2.75	4.93	nonC4-40	4.17	-	0.44	-
nonC4-14	5.61	5.39	4.25	3.94	nonC4-41	5.33	6.67	3.25	3.75
nonC4-15	6.12	6.11	4.06	3.63	nonC4-42	8.06	8.78	5.81	5.88
nonC4-16	4.06	6.83	2.38	3.69	nonC4-43	5.94	4.89	3.44	3.06
nonC4-17	5.67	6.94	5.44	4.44	nonC4-44	2.94	6.61	1.81	4.50
nonC4-18	3.44	4.00	1.81	2.56	nonC4-45	5.67	-	5.81	-
nonC4-19	5.83	6.00	7.00	5.63	nonC4-46	3.00	5.89	2.88	4.50
nonC4-20	3.89	-	2.13	-	nonC4-47	4.39	4.67	6.69	3.81
nonC4-21	3.78	6.61	1.00	5.50	nonC4-48	6.11	5.44	5.50	5.00
nonC4-22	4.44	7.28	2.25	4.56	nonC4-49	6.89	6.83	5.31	4.38
nonC4-23	7.78	-	5.38	-	nonC4-50	4.33	4.17	3.00	3.94
nonC4-24	4.59	6.33	3.19	4.13	nonC4-51	5.61	7.11	4.38	3.94
nonC4-25	4.17	6.28	2.19	2.38	nonC4-52	6.06	6.72	4.81	4.44
nonC4-26	6.50	7.11	2.38	2.44	nonC4-53	6.67	6.17	2.56	2.69
nonC4-27	6.94	7.61	5.69	6.00	**Mean**	5.41	6.58	4.15	4.63

Considering Table 14 for subscale Part 1, the mean for the control group at the beginning of the semester is 5.41 points (Some Confidence) and the mean at the end of the semester is 6.58 points (Much Confidence). This is a difference of 1.17 points. The standard deviation for Part 1 at the first administration is 1.55 and for the last administration is 1.18. The range of scores for Part 1 at the first administration is 2.44 – 8.17 (a difference of 5.73 points) and at the last administration is 4.00 – 8.94 (a difference of 4.94 points). Additionally, the mean of subscale Part 2 at the beginning of the semester is 4.15 points (Some Confidence) and the mean at the end of the semester is 4.63 points (Some Confidence). This is a difference of 0.48 points. The standard deviation for Part 2 at the first administration is 1.76 and for the last administration is

1.60. The range of scores for Part 2 at the first administration is 0.44 – 8.13 (a difference of 7.69

points) and at the last administration is 2.07 – 7.94 (a difference of 5.87 points). Comparisons

between the subscale mean scores and standard deviations of the treatment and control groups

are in Table 15 and a comparison of just the subscale mean scores is in Figure 9.

Table 15: Comparison of Treatment and Control *MSES* Subscale Scores

		Part 1		Part 2	
		MSES 1	*MSES 3*	*MSES 1*	*MSES 3*
Comparison of Means	Treatment Group	4.93	7.46	4.64	6.18
	Control Group	5.41	6.58	4.15	4.63
	Difference from perspective of treatment group	-0.48	0.88	0.49	1.55
Comparison of Standard Deviations	Treatment Group	1.21	1.00	1.56	1.58
	Control Group	1.55	1.18	1.76	1.60

Figure 9: Comparison of Treatment and Control *MSES* Subscale Scores
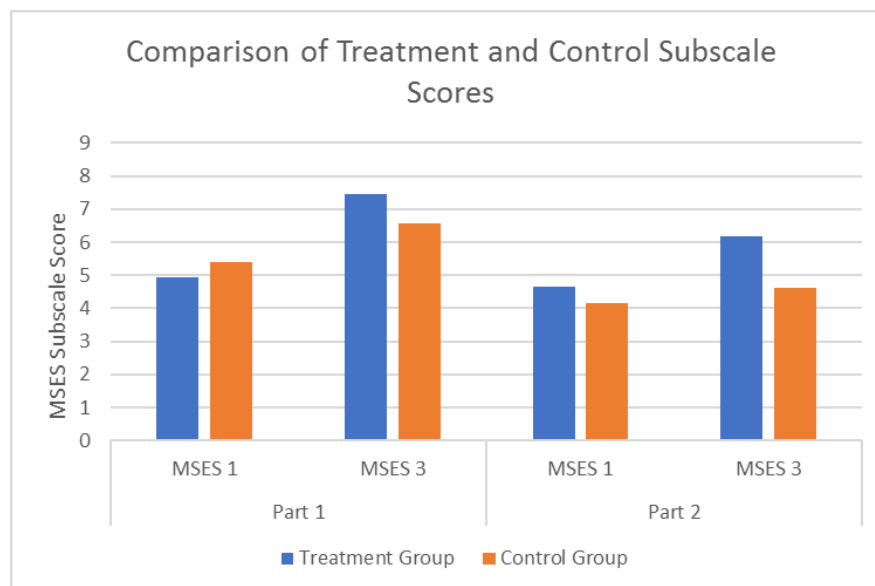
Notice, in Table 15, the control group had a mean Part 1 subscale score that was 0.48

points larger and a mean Part 2 subscale score that was 0.49 points smaller than those of the

treatment group. The Independent Samples Test (T-test) in Table 13 reports no significant

differences between the two groups at the beginning of the semester on either subscale; Part 1

subscale ($t(70) = 1.23$, $p = .223$), and Part 2 subscale ($t(70) = -1.07$, $p = .290$). To test the

differences between the two groups' mean subscale scores across time, since both groups showed

improvements, I conducted a Repeated Measures ANOVA in SPSS with a Bonferroni correction

for each subscale. This test considers the change in both subscale scores made by both groups

(from *MSES* 1 to *MSES* 3) and whether those changes are significant, ultimately considering a

time-by-treatment interaction. This is important because it is unclear whether the size of the

improvement made by both groups is similar. The tests of the interaction for both subscales were

significant, indicating that the treatment participants made greater improvements in *MSES*

subscale scores across time than those of the control group; Part 1: Wilks' Lambda = 0.44, $F(1,$

$61) = 76.35$, $p < .001$, and Part 2: Wilks' Lambda = 0.75, $F(1, 61) = 20.39$, $p < .001$. A sensible

interpretation of the statistically significant interaction results was that both groups improved

over time on both subscales, but the treatment group's scores on the two subscales improved

more even though both groups had similar starting points on the subscales of *MSES* 1 ($M_{\text{Treatment-Part1}} = 4.93$, $M_{\text{Treatment-Part2}} = 4.64$ and $M_{\text{Control-Part1}} = 5.41$, $M_{\text{Control-Part2}} = 4.15$). Therefore, the

results of the subscales on this measure (the *MSES*) implies that the levels of self-efficacy

between the two groups were different at the end of the semester; the treatment group ($M_{\text{Treatment-Part1}} = 7.46$ and $M_{\text{Treatment-Part2}} = 6.18$) ended with higher levels of self-efficacy on the subscales of

MSES 3 than the control group ($M_{\text{Control-Part1}} = 6.58$ and $M_{\text{Control-Part2}} = 4.63$).

In summary, the control and treatment groups have no statistically significant differences

between them on the first administration of the *MSES* survey in terms of their total scores and the

scores of the two subscales. At the beginning of the semester, the treatment group's average total

MSES score is 4.79 points (Some Confidence) and the control group's average total *MSES* score

is 4.81 points (Some Confidence), a difference of 0.02 points. The average subscale score for Part 1 at the beginning of the semester for the treatment group is 4.93 points (Some Confidence) and the control group is 5.41 points (Some Confidence), a difference of 0.48 points. The average subscale score for Part 2 at the beginning of the semester for the treatment group is 4.64 points (Some Confidence) and the control group is 4.15 points (Some Confidence), a difference of 0.49 points. Upon completion of the treatment course, statistically significant differences did result between the two groups based on the improvements made across time on the *MSES* in terms of their total scores and the two subscales, all in favor of the treatment group. In particular, at the end of the semester, the treatment group's average total *MSES* score is 6.86 points (Much Confidence) and the control group's average total *MSES* score is 5.66 points (Some Confidence). It was determined that the treatment group made greater gains at the end of the semester than those made by the control group, meaning that a time-by-treatment interaction was significant for the total *MSES* scores. The average subscale score for Part 1 at the end of the semester for the treatment group is 7.46 points (Much Confidence) and the control group is 6.58 points (Much Confidence). The average subscale score for Part 2 at the end of the semester for the treatment group is 6.18 points (Much Confidence) and the control group is 4.63 points (Some Confidence). It was determined that the treatment group made greater gains at the end of the semester than those made by the control group, meaning that a time-by-treatment interaction was significant for the two subscale *MSES* scores as well. These results suggests that it appears the treatment course had a significant effect on the self-efficacy of the treatment group, because even though both groups started at similar points for their total *MSES* and two subscale scores, the treatment group showed greater improvements than those made by the control group at the end of the semester.

101

Research Question 3: Changes in Self-Efficacy Based on Reports by the Participants

Research Question 3 is: "According to students' journals and interviews from the treatment group, have changes in their level of self-efficacy occurred? And, what do they think are the reasons for these changes or absence of changes?" It is important to clarify for this research question, that the treatment participants are the reporters of data, in the form of Journals and Interviews. These Journals and Interviews were then analyzed, as described in Chapter 3, by the researcher to determine if changes in self-efficacy occurred. To determine whether a participant experienced an increase, decrease, or no change in their self-efficacy and answer the first part of this research question, I analyzed their ratings on Journals 1 and 5 and Interviews 1 and 2. Recall that the ratings are assigned based on the percentage of positive and negative statements that were coded in the Journals and Interviews, see Table 16.

positive statements + negative statements = 100%

Table 16: Rating Scale of Journals and Interviews

Rating Category	Description	Percentages of Statements
Very Negative Self-efficacy	all negative, no positive statements OR overwhelmingly negative	85%-100% of the statements are negative
Moderately Negative Self-efficacy	many more negative than positive statements	75%-84% of the statements are negative
Somewhat Negative Self-efficacy	a few more negative than positive statements	61%-74% of the statements are negative
Neutral Self-Efficacy	fairly equal negative and positive statements	50%-60% of the statements are positive and/or 50%-60% are negative
Somewhat Positive Self-efficacy	a few more positive than negative statements	61%-74% of the statements are positive
Moderately Positive Self-efficacy	many more positive than negative statements	75%-84% of the statements are positive
Very Positive Self-efficacy	all positive, no negative statements OR overwhelmingly positive	85%-100% of the statements are positive

Further, I paired colors with the ratings in order to clearly illustrate any increases, decreases, or no change in self-efficacy experienced by participants in a visual manner, see Figure 10.

Figure 10: Rating Scale of Journals and Interviews with Colors

Very Negative	Moderately Negative	Somewhat Negative	Neutral	Somewhat Positive	Moderately Positive	Very Positive

The results of the data analysis phase for Journals and Interviews are explicitly shown in Table 17. Each entry in the table corresponds to a Journal or Interview and indicates the rating on each measure for each participant along with the numerical score that was used to determine the rating based on Table 16. This numerical score shows the number and type of statements made by participants in the form of (negative – positive – variable). The first row contains a note about which week during the semester each measure was administered to provide perspective to the reader.

Table 17: Results of Journal and Interview Data Analysis

Participant	Journals					Interviews	
	1 Week 1	2 Week 5	3 Week 10	4 Week 14	5 Week 15	1 Weeks 2-3	2 Weeks 14-15
1	Somewhat Negative (5-2-1)	Neutral (1-1-2)	Very Positive (1-6-0)	Very Positive (0-7-0)	Moderately Positive (1-5-2)	Neutral (8-11-5)	Moderately Positive (8-32-1)
2	Very Positive (0-6-0)	Very Positive (0-3-1)	Moderately Positive (1-3-0)	Neutral (3-4-1)	Moderately Negative (3-1-1)	Very Positive (8-50-7)	Very Positive (5-31-5)
3	Neutral (2-3-1)	Neutral (2-3-1)	Neutral (4-3-0)	Somewhat Negative (2-1-1)	Neutral (2-2-2)	Somewhat Negative (22-11-6)	Neutral (22-24-3)
4	Somewhat Positive (2-4-1)	Moderately Positive (2-6-1)	Moderately Positive (1-5-2)	Neutral (4-3-1)	Neutral (3-4-1)	Neutral (22-28-6)	Somewhat Positive (16-32-6)
5	Neutral (4-6-0)	Moderately Positive (1-5-1)	Very Positive (0-9-0)	Very Positive (0-8-0)	Very Positive (1-6-0)	Neutral (10-10-3)	Very Positive (3-29-2)
6	Moderately Positive (1-3-0)	Moderately Positive (1-5-0)	Moderately Positive (1-5-2)	Moderately Positive (1-5-1)	Moderately Positive (2-6-1)	Neutral (20-16-5)	Somewhat Positive (9-23-4)

Table 17 (cont'd)

7	Neutral (4-4-0)	Moderately Positive (1-5-0)	Neutral (4-6-1)	Somewhat Negative (4-2-2)	Neutral (3-2-0)	Moderately Positive (6-26-4)	Somewhat Positive (13-34-3)
8	Moderately Negative (5-1-0)	Somewhat Negative (5-3-0)	Neutral (2-3-2)	Moderately Negative (4-1-2)	Somewhat Positive (2-4-1)	Somewhat Negative (18-7-2)	Neutral (15-21-2)
9	Somewhat Negative (5-3-0)	Somewhat Positive (2-4-0)	Moderately Positive (1-3-2)	Neutral (3-4-2)	Somewhat Positive (3-6-1)	Somewhat Negative (13-8-6)	Moderately Positive (5-15-2)
10	Neutral (2-2-0)	Very Positive (0-4-1)	Very Positive (1-6-0)	Very Positive (0-7-0)	Very Positive (0-3-0)	Somewhat Positive (7-13-2)	Very Positive (1-27-0)
11	Neutral (3-3-0)	Very Positive (0-6-0)	Somewhat Positive (2-4-0)	Very Positive (0-6-0)	Very Positive (0-6-1)	Moderately Positive (11-3-2)	Very Positive (1-20-4)
12	Moderately Negative (3-1-0)	Somewhat Positive (2-5-2)	Somewhat Positive (3-6-2)	Moderately Positive (1-4-4)	Moderately Positive (1-5-3)	Somewhat Negative (18-9-2)	Somewhat Positive (13-27-8)
13	Moderately Negative (5-1-0)	Very Positive (1-7-0)	Somewhat Positive (3-7-1)	Very Positive (1-6-1)	Very Positive (1-7-1)	Neutral (21-14-3)	Somewhat Positive (7-20-11)
14	Very Negative (8-1-0)	Neutral (4-4-0)	Neutral (3-4-1)	Moderately Negative (4-1-2)	Somewhat Negative (5-2-1)	Moderately Negative (19-5-2)	Moderately Negative (32-9-6)
15	Neutral (3-3-0)	Moderately Negative (3-1-1)	Moderately Positive (1-5-1)	Somewhat Positive (2-4-2)	Somewhat Negative (4-2-0)	Neutral (15-14-7)	Neutral (14-16-8)
16	Moderately Negative (3-1-0)	Neutral (2-3-1)	Somewhat Positive (3-5-2)	Neutral (4-4-0)	Neutral (5-5-2)	Moderately Negative (27-7-3)	Neutral (35-23-7)
17	Neutral (2-2-0)	Somewhat Positive (1-2-0)	Very Positive (0-5-1)	Very Negative (3-0-2)	Somewhat positive (1-2-0)	Moderately Positive (5-15-0)	Neutral (6-9-2)
18	Somewhat Negative (5-2-0)	Somewhat Positive (2-4-0)	Somewhat Positive (1-2-0)	Moderately Positive (1-4-1)	Moderately Positive (1-4-1)	Somewhat Negative (17-7-2)	Somewhat Positive (10-21-10)
19	Neutral (3-2-1)	Neutral (2-3-0)	Somewhat Positive (3-5-0)	Moderately Positive (2-6-0)	Very Positive (1-6-1)	Moderately Negative (21-6-3)	Somewhat Positive (15-35-5)

Journals. Considering the Journals at the start (Journal 1) and end (Journal 5) of the

semester: 13 students showed an increase in self-efficacy, three students showed no change, and

three students showed a decrease in their self-efficacy. There was evident fluctuation in self-

efficacy ratings for many participants between these two time points and these will be discussed

later in this section. Participant 13 showed the greatest increase in self-efficacy, jumping from

Moderately Negative at the beginning of the semester to Very Positive at the end of the semester, a change across five levels in Figure 10. Similarly, Participant 12 made a gain of four levels, from Moderately Negative at the beginning of the semester to Moderately Positive at the end of the semester. Participant 2 showed the greatest decrease in self-efficacy in his Journals, declining from Very Positive at the beginning of the semester to a Moderately Negative rating at the end of the semester, a change across five levels. A visual representation that shows each Journal administration and the number of participant responses that are in each category from the rating scale in Table 16, is shown in Figure 11. This representation makes it easier to see the treatment group's shift from more responses at a Neutral rating or below to more responses above Neutral as the semester progresses. In particular, for Journal 1, there are eight participants with some negative rating, eight participants with a neutral rating, and three participants with some positive rating. This shifts immediately with Journal 2, as only two participants have some negative rating, five participants have a neutral rating, and 12 participants have some positive rating. The results of the remaining Journals support this shift and maintain greater numbers of positive ratings than negative and neutral ratings combined.

Figure 11: Self-Efficacy Ratings from Journal Responses
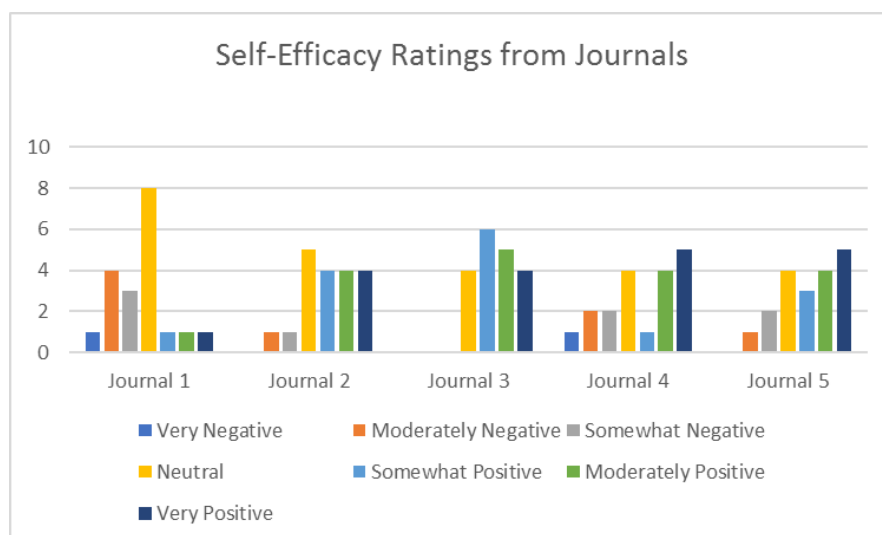
Table 18 illustrates the number of levels gained (+) or lost (-) from Journal 1 to Journal 5. The last row illustrates the average gain for the entire group, which is a gain of about 1.5 levels. This shows that a student who enters the treatment course with a Somewhat Negative rating could be expected to make a rise of 1.5 levels and end the semester with a Somewhat Positive rating, similar to the experience of Participant 9.

Table 18: Number of Levels Gained or Lost According to Analysis of Participant Journals 1 to 5

Participant	Number of Levels Gained or Lost
1	+3
2	-5
3	0
4	-1
5	+3
6	0
7	0
8	+3
9	+2
10	+3
11	+3
12	+4
13	+5
14	+2
15	-1
16	+2
17	+1
18	+3
19	+3
Average	1.58

Interviews. For the Interviews, 14 participants showed an increase in self-efficacy, three participants showed no change, and two participants showed a decrease in their self-efficacy. Participants 5, 9, and 19 showed the greatest gains in their Interviews, each of them gaining three levels in their self-efficacy for mathematics. The only participants whose self-efficacy decreased, according to their Interviews, were Participant 7 with a drop of one level and Participant 17 with

a drop of two levels. A visual representation that shows each Interview administration and the number of participant responses that are in each category from the rating scale in Table 16, is shown in Figure 12. This representation makes it is easier to see the treatment group's shift from more responses at a Neutral or below rating to more responses above Neutral as the semester progresses. In particular, for Interview 1, there are eight participants with some negative rating, six participants with a neutral rating, and five participants with some positive rating. A noticeable shift is present with the results of Interview 2, as only one participant has some negative rating, five participants have a neutral rating, and 13 participants have some positive rating.

Figure 12: Self-Efficacy Ratings from Interview Responses
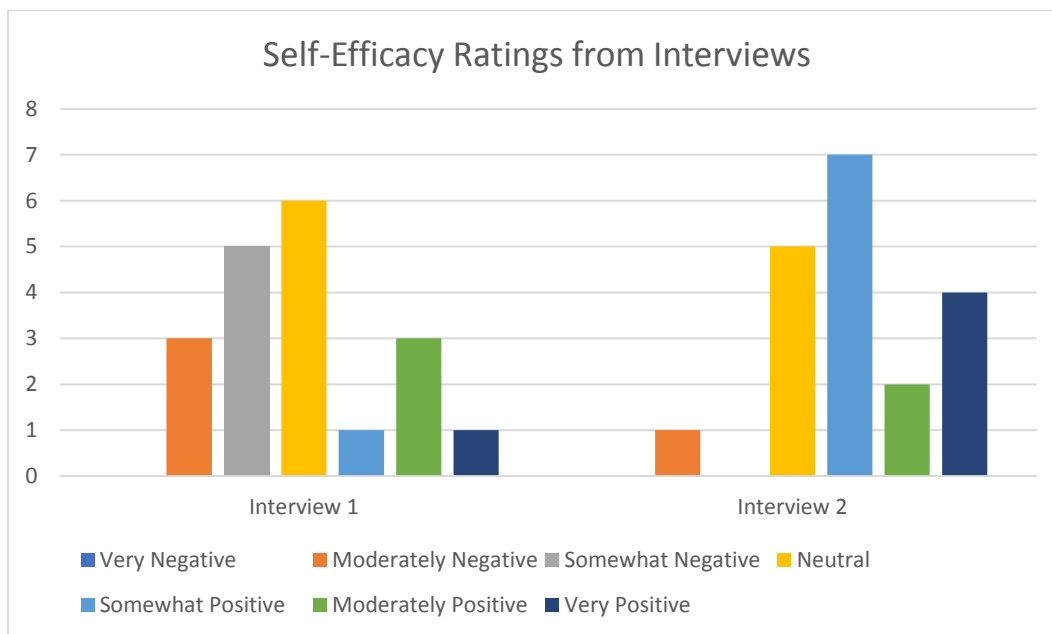
Table 19 illustrates the number of levels gained (+) or lost (-) from Interview 1 to Interview 2. The last row illustrates the average gain for the entire group, which is a gain of about one level. This shows that a student who enters the treatment course with a Neutral rating

could be expected to make a rise of 1 level and end the semester with a Somewhat Positive

rating, similar to the experience of Participant 4.

Table 19: Number of Levels Gained or Lost According to Analysis of Participant Interviews

Participant	Number of Levels Gained or Lost
1	+2
2	0
3	+1
4	+1
5	+3
6	+1
7	-1
8	+1
9	+3
10	+2
11	+1
12	+2
13	+1
14	0
15	0
16	+2
17	-2
18	+2
19	+3
Average	1.16

Journals and interviews. When the analysis results of the Journals and Interviews are

viewed together (Table 20), 11 participants showed gains in self-efficacy for both their Journals

and Interviews from the beginning to end of semester, with some fluctuations throughout.

Additionally, no students showed no change on both measures and no students' self-efficacy

decreased on both measures. Participants 4 and 17 showed a combination of increase on one

measure and decrease on the other measure, Participants 3, 6, 14 showed a combination of

increase on one measure and no change on the other measure, and Participants 2, 7, and 15

showed a combination of decrease on one measure and no change on the other measure. Table *20* illustrates the combinations for each of the treatment participants.

Table 20: Comparison of Journal and Interview Results

Participant	Journals	Interviews
1	Increase	Increase
2	Decrease	No Change
3	No Change	Increase
4	Decrease	Increase
5	Increase	Increase
6	No Change	Increase
7	No Change	Decrease
8	Increase	Increase
9	Increase	Increase
10	Increase	Increase
11	Increase	Increase
12	Increase	Increase
13	Increase	Increase
14	Increase	No Change
15	Decrease	No Change
16	Increase	Increase
17	Increase	Decrease
18	Increase	Increase
19	Increase	Increase

I will discuss each of these four combinations and the participants in those groupings in the following paragraphs. Quotes from the participants will be used in order to begin to answer the second part of the research question about what reasons they attribute to these changes or absence of changes in self-efficacy. More discussion about these reasons will follow once the results of self-efficacy are presented for the first part of this research question. In the following paragraphs, a numerical score for each participant's Journal or Interview will be given next to its rating. The numerical score, assigned during the data analysis phase, is composed of the number of statement types made by the participant in the form of (negative – positive – variable), and informed the rating, which are seen in Table 17.

Increase in both measures. Eleven participants showed increase in both measures: 1, 5, 8, 9, 10, 11, 12, 13, 16, 18, and 19. Since so many students showed an increase on both measures, I will to highlight the specific stories of Participants 13, 18, and 19. These three participants were chosen because they each had a different base level of self-efficacy according to analysis of their Journals and Interviews, and each of these starting levels were similar to those of their classmates in this grouping. But, first, I will provide, in Table 21, each of these eleven participants' responses to two questions from Interview 1 to Interview 2 to provide evidence for their positive shifts.

Table 21: Participant Responses on Two Interview Questions Across Time

Participant	"How do you currently feel about math?"		"Do you feel you have been successful with math?"	
	Interview 1 Response	Interview 2 Response	Interview 1 Response	Interview 2 Response
1	Missing *Question was skipped over by student interviewer	"Really good, my math class has been so helpful. I feel way more confident in myself"	"Not too bad. I mean I had my ups and downs"	"For sure. My math grade has been as high as it's ever been"
5	"Right now I think I'm pretty okay. I just…some things do get to me, so when I stress…when I overcomplicate a problem or a word problem or anything, I just…I feel like I can't do it. But then when I get help I calm down"	"I've always liked math, so it really hasn't been a tough subject for me because I, knowing that there's an exact answer. I feel really confident in math so I'm not really struggling"	"Yeah, I do. I feel like I've done pretty good in math"	"Yeah, I do"
8	"Still don't like it…. Something that doesn't come easy. And if it's basic like adding, subtracting, multiplying, dividing, that's easy. But when it starts to get to fractions, and all that, it's no…. Way too complicated"	"A lot better than I did at the beginning of the semester because it's a lot easier now"	"No"	"Oh yeah. I got a 94% in that class rather than a 58% like I did in high school for four years"
9	"I didn't feel confident at all"	"I feel more confident and can actually ask for help…. And not feel shameful of it"	"For the most part. Freshman year, not so much"	"I would say I've been more successful"
10	"I was always good at math. So…I feel pretty confident"	"Currently I feel really good in math. Math is my second highest grade right now this semester…. Math has just always came easy to me and I feel really confident going into a test without studying. But I will study, but I feel confident, just comes second nature to me"	"Yeah"	"Yes…. I know the material enough where I can teach somebody else"
11	"So, that's probably been my biggest struggle probably. And all the courses. Math really gets me"	"Pretty good. I honestly like being in this class is more one on one I would say than in high school like teaching through the steps a little bit easier"	"No"	"Yeah, I think I've been successful with math"

Table 21 (cont'd)

12	"Right now, during our program, I'm getting a lot of help. Lauren's doing really good at explaining things, and I feel like I'm learning a lot more" *Prior to college was not clarified by student interviewer	"I actually feel really good compared to how I started. I thought I was going to be super bad at it but I've actually. really improved"	"No…. I've always struggled with just simple stuff like adding in my head or subtracting big numbers. I just have to do writing it down on a piece of paper or something. It doesn't come quick to me, I guess"	"Yeah"
13	"Math is probably not my strongest subject. It was probably the worst subject. I struggle so much in math, but I don't know. I try to ask a lot of people for help"	"I actually like math now. I used to not like it. I have a positive mindset going into class instead of dreading it every day"	"I'd say I was okay at it. I wouldn't say I was really good at it. Probably could improve"	"Yes, I do feel that way"
16	"Still unsure, I understand some, I understand more than I did, but I'm still not comfortable with it. I wouldn't be able to help other kids if they needed help"	"It's not my favorite subject, but I feel more confident about it. It makes me feel better when I know that I do something right or like I understand something"	"No"	"Not like through and through, I've struggled but from like the beginning until now, I think I have improved but I wasn't successful all the time"
18	"Not that confident…. Math hasn't always been my strongest subject. Math is one of my weakest. I usually don't feel as confident when I do math"	"I'm feeling better about math. Like the logarithm stuff, that had gotten a lot easier since…I've been going to office hours and stuff. There's still some…little parts that I struggle with…but other than that…I'm feeling like more confident about math now"	"Kind of"	"Yeah, I think I've been. Well, I think I've been more successful at math than I have when I first started"
19	"…. I don't know, sometimes I'm really not good at math. I know I can solve for *x* and stuff, but when it comes to linear equations on a graph or factoring a fraction or something, I know I'm not good at that. I have to have steps in front of me to know how to solve it"	"I feel actually okay with it. Somebody can teach me something once and then I'll understand it. I get maybe a little tricked up here and there, but now I feel okay. I'm not really scared to go into my Chemistry course next year. I mean next semester. Yeah, I feel pretty confident with it now."	"Not really…. I don't think I was going to be successful and pass any of my math classes that I had"	"Yes, I have"

The responses in Table 21 help to illustrate how students' self-efficacy appears to have increased upon completion of the treatment course. For example, the positive shift is evident by comparing the two answers that Participant 9 gives to the question, "How do you currently feel about math?" In Interview 1 she said, "I didn't feel confident at all." This can be directly compared to her response in Interview 2 when she said, "I feel more confident and can actually ask for help.... And not feel shameful of it." Similar comparison can be made for other participants. Let us now consider the stories of Participants 13, 18, and 19.

The analysis of Participant 13's Journals show she began the semester with a Moderately Negative (Journal 1: 5-1-0) rating. Her entries appeared to increase significantly in the first half of the semester to a Very Positive rating (Journal 2: 1-7-0) but declined slightly to Somewhat Positive (Journal 3: 3-7-1) at the middle of the semester. She, then, spent the remainder of the semester with a Very Positive rating (Journal 4: 1-6-1 and Journal 5: 1-7-1). The analysis of her Interviews suggest a similar story, with her increasing from Neutral (Interview 1: 21-14-3) to Somewhat Positive (Interview 2: 7-20-11) during the semester. To illustrate this shift with evidence from her Interviews, her response to, "Do you feel you have been successful with math?" was, "I'd say I was okay at it. I wouldn't say I was really good at it. Probably could improve" (Interview 1) compared to, "Yes, I do feel that way" (Interview 2).

The analysis of Participant 18's Journals show he began the semester with a Somewhat Negative (Journal 1: 5-2-0) rating. His comments increased to Somewhat Positive (Journal 2: 2-4-0 and Journal 3: 1-2-0) and toward the end of the semester increased again to Moderately Positive (Journal 4: 1-4-1 and Journal 5: 1-4-1). The analysis of his Interviews tell a similar story and he appears to show gain from Somewhat Negative (Interview 1: 17-7-2) to Somewhat Positive (Interview 2: 10-21-10) during the semester. To illustrate this shift with evidence from

113

his Interviews, his response to, "How do you currently feel about math?" was, "Not that confident…. Math hasn't always been my strongest subject. Math is one of my weakest. I usually don't feel as confident when I do math" (Interview 1) compared to his response in Interview 2:

I'm feeling better about math. Like the logarithm stuff, that had gotten a lot easier since…I've been going to office hours and stuff. There's still some…little parts that I struggle with…but other than that…I'm feeling like more confident about math now.

The analysis of Participant 19's Journals show she began the semester with a Neutral (Journal 1: 3-2-1 and Journal 2: 2-3-0) rating, which continued through Exam 1. Before Exam 2, her comments changed to Somewhat Positive (Journal 3: 3-5-0). She then increased to Moderately Positive (Journal 4: 2-6-0) around the time of Exam 3 and ended the semester with a Very Positive (Journal 5: 1-6-1) rating. The analysis of her Interviews tell a similar story and show her increasing from Moderately Negative (Interview 1: 21-6-3) at the beginning of the semester to a Somewhat Positive (Interview 2: 15-35-5) rating. This change can be illustrated by her responses to the question, "Do you feel that you have been successful with math?". In Interview 1 she said, "Not really…. I don't think I was going to be successful and pass any of my math classes that I had." In Interview 2 her response changed to a confident sounding, "Yes, I have". She noted an example of feeling successful later in the Interview and said, "I really understand [logarithms]. I was like, wow, this is easy. I don't know why I didn't know how to do it [in high school]".

Increase and decrease combination. Two participants showed an increase on one measure and a decrease on the other measure. This happened for Participants 4 and 17.

The analysis of Participant 4's Journals show she began the semester with a Somewhat Positive (Journal 1: 2-4-0) rating, then increased to Moderately Positive (Journal 2: 2-6-1 and

Journal 3: 1-5-2) for some time during the middle of the semester. This upward shift is supported by many positive statements that she made. An example was when she spoke about her Exam 1 score and said it, "made me feel very confident about myself and how I feel about math" (Journal 3). She then decreased to a Neutral (Journal 4: 4-3-1 and Journal 5: 3-4-1) rating for the end of the semester, which is supported by a combination of statements like, "I scored lower [on Exam 3] than I usually would, but this was because I struggled very hard with logarithms…My experience in this math course this semester was actually awesome" (Journal 5). To contrast, the analysis of her Interviews suggest that she began the semester with a Neutral (Interview 1: 22-28-6) rating and then increased to Somewhat Positive (Interview 2: 16-32-6) by the end of the semester. Neutral and Somewhat Positive are only one level apart, but in fact her Journals did show some gain and then loss. In this case, the analysis of the Journals presented a different story than that of the Interviews, which showed a slight gain in self-efficacy.

The analysis of Participant 17's Journals show he began the semester with a Neutral (Journal 1: 2-2-0) rating, then increased to Somewhat Positive (Journal 2: 1-2-0) and further to Very Positive (Journal 3: 0-5-1) at the middle of the semester. This upward shift is supported by many positive statements including, "I feel confident about being ready for [Exam 1]," (Journal 2) and, "I have received numerous successes on tests, quizzes, and homework too" (Journal 3). His self-efficacy appeared to take a drastic turn and he fell to Very Negative (Journal 4: 3-0-2) between Exams 2 and 3. Evidence for this is given with statements like, "I felt bad about my Exam 2 score…I was upset with myself about it" (Journal 4). However, he was able to recover and end the semester with a Somewhat Positive (Journal 5: 1-2-0) rating, which is one level greater than he initially began the semester. The analysis of his Interviews tells a different story about a loss in self-efficacy instead of a gain. Analysis of the Interviews suggest he began with a

115

Moderately Positive (Interview 1: 5-15-0) rating and then he fell to a Neutral (Interview 2: 6-9-2) rating by the end of the semester. This downward shift is illustrated by his response to the question, "Have your feelings about math changed over time this semester?" He said, "Yeah, just a little…. Throughout high school, math was easy for me, because I knew the stuff. Now I don't know it, because it was back from Algebra 1" (Interview 2). Although, he still indicated that he felt successful in math this semester and had good feelings about it. The combination of statements like these illustrate the Neutral nature of Interview 2 and the appearance of decline in self-efficacy.

    Increase and no change combination. Three participants showed an increase on one measure and no change on the other measure. This happened for Participants 3, 6, and 14.

    The analysis of Participant 3's Journals show she began the semester with a Neutral (Journal 1: 2-3-1) rating and maintained this level of self-efficacy for the first half of the semester (Journal 2: 2-3-1 and Journal 3: 4-3-0). However, between Exams 2 and 3, she decreased slightly to Somewhat Negative (Journal 4: 2-1-1), attributing this to, "getting a bad grade [on Exam 2 which] made me feel awful about myself" (Journal 4). She then gained, "more self-confidence in math" (Journal 5) and increased back to Neutral (Journal 5: 2-2-2) at the end of the semester. The analysis of her Interviews, however, suggest a similar story but with a slight increase from the beginning of the semester with a Somewhat Negative (Interview 1: 22-11-6) rating to the end of the semester with a Neutral (Interview 2: 22-24-3) rating. This increase is illustrated by her response to the question, "Do you think your feelings about math have changed over this semester?" She said, "Yeah, I just thought coming into math, I wasn't going to know anything and struggle. But coming [to] the end of the semester, I really think I did really well in math" (Interview 2). Some hesitation toward mathematics is still present throughout the

Interview, especially in relation to the struggle she associates with fractions and word problems. The combination of these types of statements supports the assigned Neutral rating.

The analysis of Participant 6's Journals are consistent with a Moderately Positive rating for the entire semester (Journal 1: 1-3-0, Journal 2: 1-5-0, Journal 3: 1-5-2, Journal 4: 1-5-1, and Journal 5: 2-6-1). The analysis of his Interviews suggests, however, that he began the semester at a Neutral (Interview 1: 20-16-5) rating and ended the semester at a Somewhat Positive (Interview 2: 9-23-4) rating. This positive shift is supported by his response to, "Describe how you currently feel about math." He said, "I currently feel really good about my math. I feel I've improved and it's coming to me a lot easier" (Interview 2). Later in the Interview he also says, "But now I'm thinking more positive. I'm understanding and grasping it more" when comparing his experience of receiving assignments in high school to the current math course (Interview 2). These statements help to illustrate the positive shift that was identified by analysis.

The analysis of Participant 14's Journals show she began the semester with a Very Negative (8-1-0) rating, then increased to Neutral (Journal 2: 4-4-0 and Journal 3: 3-4-1) during the middle of the semester. Then, between Exams 2 and 3, she decreased noticeably to Moderately Negative (Journal 4: 4-1-2), attributing this decline to her poor Exam 2 score. She said "It made me feel upset and feel sad. Made me feel like I am bad at math" (Journal 4). She finished the semester slightly higher at Somewhat Negative (Journal 5: 5-2-1). Her reflection on the semester indicates:

> I feel better because when I struggled before [during high school,] I was failing and I am
> not failing now when I struggle. I have felt some succes [*sic*] when I am understanding.
> When I get to the test I loose cofidence [*sic*] in myself. (Journal 5)

The analysis of her Interviews suggest that she remained at Moderately Negative for the entire semester (Interview 1: 19-5-2 and Interview 2: 32-9-6). To illustrate this, her response to "Have your feelings about math changed over time?" in Interview 1 was, "No. I've always not liked math…. I feel that I need a lot of work in math and that I don't do well in math and that it really upsets me sometimes." When asked the same question, but about the current semester, in Interview 2, she said, "Yes, because in the beginning I was very confident with myself because we started off with a little bit more easier stuff and as it got harder, my confidence went down with it." These statements illustrate similar feelings at the beginning of the semester and at the end of the semester. The Interview 2 response even suggests that she felt an increase in her confidence level toward the beginning of the semester, but then lost that gain and remained at a Moderately Negative rating at the end of the semester.

Decrease and no change combination. Three participants showed a combination of decrease on one measure and no change on the other measure. This happened for Participants 2, 7, and 15.

The analysis of Participant 2's Journals show he began the semester with a Very Positive (Journal 1: 0-6-0) rating, which he maintained until around the middle of the semester (Journal 2: 0-3-1). He wrote, "When it comes to Math, I am alway [*sic*] very thrilled and excited about it….I hope to develop my skills even more this year and hopefully be the best in class," in his Journal 1 (Math Autobiography). He then slightly decreased to Moderately Positive (Journal 3: 1-3-0) around the time of Exam 2. Then, while being directly involved in significant classroom drama with other classmates, he appeared to significantly decrease to Neutral (Journal 4: 3-4-1) between Exams 2 and 3. His reflection on receiving his Exam 2 score notes, "It made me feel unsecure about math. I'm good at math and I know I am" (Journal 4). He decreased even more to

118

end the semester at Moderately Negative (Journal 5: 3-1-1). Reflecting on the semester shows that he observed a negative change: "This semester with math has been frustrating. I started really good and have slowy [*sic*] gotten worse" (Journal 5). The analysis of his Interviews, however, suggest that he remained Very Positive throughout the entire semester (Interview 1: 8-50-7 and Interview 2: 5-31-5). His response to "Thinking about this semester, do you feel that you've been successful with math?" was:

> Yeah, I absolutely do. I know that I may not have the grade that I initially started with at math, but I'm not a perfect student. Failure is what motivates me to keep on going, and it drives me to study more or work harder or ask more questions. (Interview 2)

This statement helps to illustrate that his Interviews remained at a Very Positive level throughout the semester, which is an interesting contrast to his Journals and will be discussed further in Research Question 4. Further, he is the only participant to begin the treatment at Very Positive for either measure, which was unexpected given the characteristics that typical C4 Scholar students begin with. Additionally, he is the only student to begin the treatment with any level of positive self-efficacy on both measures. There are six other participants that began at a positive level of self-efficacy on one of the measures, but not for both Journals and Interviews.

The analysis of Participant 7's Journals show she began the semester with a Neutral (Journal 1: 4-4-0) rating and then noticeably increased to Moderately Positive (Journal 2: 1-5-0) just before Exam 1 was given. She, then, around the middle of the semester was diagnosed with an illness that would remove her from the classroom for almost a week and she fell back to a Neutral (Journal 3: 4-6-1) rating before Exam 2 was given. She referenced how this affected her in Journal 3, without explicitly mentioning the illness, by having said, "One issue I'm having was in week 7. I was confused on most of it and didn't fully understand it." She was absent for

119

three out of four class periods that week because of the illness. Between Exams 2 and 3 she fell slightly to Somewhat Negative (Journal 4: 4-2-2). In Journal 4, she was very concerned about how her Exam 2 score was going to affect her in the course. She said the score "made me feel as if I'm not gonna do good on the rest of my test or Quiz 5. I also am worried that I'm gonna fail my class because of this one Exam" (Journal 4). She was able to come back to a Neutral (Journal 5: 3-2-0) rating at the end of the semester, which is supported by the following statement about her Exam 3 score, "It was an improvement from my last Exam which I was happy about but it still made me feel defeaded [*sic*]." The analysis of her Interviews, however, suggest that she began the semester at Moderately Positive (Interview 1: 6-26-4) and then slightly decreased to Somewhat Positive (13-34-3).

The analysis of Participant 15's Journals show he began the semester with a Neutral (Journal 1: 3-3-0) rating, and just before Exam 1 decreased to Moderately Negative (Journal 2: 3-1-1), which he attributed to his struggles with fractions and word problems. He, then, appeared to make a significant gain to Moderately Positive (Journal 3: 1-5-1) just before Exam 2, in part because he "felt good about [his] grade on the first exam" and he received "results [he] expected" (Journal 3). Between Exams 2 and 3, he decreased slightly to Somewhat Positive (Journal 4: 2-4-2) but then decreased more to Somewhat Negative (Journal 5: 4-2-0) at the end of the semester. The evidence that supports this decrease is that he "felt terrible about [his] math Exam 3…It made me feel that I am bad at math" (Journal 5). The analysis of his Interviews suggest that he remained Neutral throughout the entire semester. This Neutral rating, that was maintained throughout the semester, is supported by his response to, "So thinking about this semester, do you feel that you've been successful with math?" He said, "Somewhat successful. Some units I have been, yes, some units, no" (Interview 2). In addition, when asked, "Have your

feeling about math changed over time this semester?", he said, "Yes. I have had my ups and downs, but I'm just happy to get it over with" (Interview 2). Statements like these helps to illustrate that he appears to have a mixture of feelings and experiences from the semester, which, based on analysis, resulted in a Neutral rating.

Participants and the variable code. In Chapter 3, I described how the positive and negative codes guided the ratings that were assigned to Journals and Interviews for each participant. As indicated in the data analysis part of that chapter, variable codes were not included in the calculation but were instead used to indicate when and how participants recognize change or variation in their experience with mathematics. Examples of variable statements given by participants are in Table 22. These examples show that a variation in experience can be recognized by students for as small as formula use all the way up to the domain of mathematics.

Table 22: Examples of Variable Statements

Data Source and Participant	Quote	Variability in Reference to
Journal 1 Participant 3	"Math is alright. It depends on what subject we are doing."	The domain
Journal 2 Participant 16	"I think that it hasn't been extremely difficult work, but it hasn't been the easiest either."	Content in Unit 1
Journal 3 Participant 6	"I struggled with the formulas at first but after seeing and studying them I felt fine."	Formula use
Journal 4 Participant 13	"Unit 3 was harder, but with some help, I understand it now."	Content in Unit 3
Journal 5 Participant 1	"When I saw what I did wrong I realized they were easy to learn from mistakes but still had enough understanding of how to do it."	An assessment
Interview 1 Participant 12	"I feel like I've improved over the past couple of years. But, I still struggled with [math]."	The domain
Interview 2 Participant 18	"I know how to set them up, but sometimes I have difficulty calculating them." [He is referring to using formulas, in general]	Formula use

These statements are reports of variation in experience with mathematics. The purpose of showing this data is to provide more details about the individual participant experience. These types of statements are important to note because they illustrate the unique journey taken by each participant, no matter the length of the journey or whether the journey ends with a positive or negative result. Any two participants may have ended the treatment with the same level of self-efficacy, within this treatment group there are many examples, but their journeys to get there can look quite different. Table 23 indicates the number of variable statements made by each participant for each Journal and Interview.

Table 23: Variable Statements from Each Participant in Journals and Interviews

Participant	Journals					Interviews	
	1	2	3	4	5	1	2
	Week 1	Week 5	Week 10	Week 14	Week 15	Weeks 2-3	Weeks 14-15
1	1	2	0	0	2	5	1
2	0	1	0	1	1	7	5
3	1	1	0	1	2	6	3
4	1	1	2	1	1	6	6
5	0	1	0	0	0	3	2
6	0	0	2	1	1	5	4
7	0	0	1	2	0	4	3
8	0	0	2	2	1	2	2
9	0	0	2	2	1	6	2
10	0	1	0	0	0	2	0
11	0	0	0	0	1	2	4
12	0	2	2	4	3	2	8
13	0	0	1	1	1	3	11
14	0	0	1	2	1	2	6
15	0	1	1	2	0	7	8
16	0	1	2	0	2	3	7
17	0	0	1	2	0	0	2
18	0	0	0	1	1	2	10
19	1	0	0	0	1	3	5
Total	**4**	**11**	**17**	**22**	**19**	**70**	**89**
Percent of Total Coded	**3.3%**	**9.4%**	**11.8%**	**15.9%**	**14%**	**11.4%**	**11.6%**

It is clear from Table 23 that, on average, the number of times that participants recognize change or variation in their experience for mathematics increases over time. What is also evident is the

greater frequency of these statements in Interviews than Journals. However, according to the last row in Table 23, when comparing the percentage of variable statements to the total number of statements coded, the two data sources show comparable results. It is interesting that an increase in the number of variable statements coincides with a general increase in self-efficacy, as indicated by the analysis of the Journals and Interviews.

Reasons According to Participants. I have presented how the analysis of Journals and Interviews indicates that participants' self-efficacy has changed upon completion of the treatment, which addresses the first part of this research question. To answer the second part of this research question, I will now focus on what the participants reported in their Journals and Interviews and that appear to coincide with these observed changes. It is important to note that when the treatment group reflected on their experiences, elements of the course and program emerged in the data for this study. Three of these elements, which were intentionally built into the course and program, were recognized by the students.

Analysis of the data, as described in Chapter 3, shows that the most commonly discussed topics in the Journals and Interviews, that appeared to affect participants, included the course instructor, support from peers in the learning community, general components of the C4 Scholar Program, support from people outside the course, and outside factors such as class social drama. However, the topic that dominated their discussions were occurrences of success and/or struggle with mathematics, much of which did not focus on grades. All 19 participants wrote or spoke about their grades in their Journals and Interviews, but on a much larger scale all 19 participants discussed specific instances of success and struggle that appeared to affect them. Further, what appeared to be the next most prominent topics were the course instructor, followed by peer support in the learning community. Support from people outside the course, components of the

C4 Scholar Program, and outside factors were also discussed by participants, but on a smaller

scale. Table 24 illustrates the number of participants and the frequency for each of the six topics.

Table 24: Commonly Reported Topics from Participant Journals and Interviews

Topic		Number of Participants		Frequency of Occurrence	
		Journals	Interviews	Journals	Interviews
Success and/or Struggle	Reference to understanding	19	19	232	137
	Reference to grades	19	17	69	52
Course Instructor		13	17	29	64
Peer Support in the Learning Community		3	17	4	59
Support from Others Outside the Course		6	16	7	41
Component of C4 Scholar Program		9	16	12	34
Outside Factors (i.e. class drama)		1	8	1	13
Total				**354**	**400**

Success and struggle. First, occurrences of success and/or struggle were discussed

frequently throughout all 19 participants' Journals and Interviews. According to the data,

students spoke or wrote about success primarily as understanding a mathematics concept and,

less frequently but still evident, earning a grade they were happy with (e.g. A or B) on an

assessment. Similarly, students spoke or wrote about struggle primarily as not understanding or

experiencing difficulties with a mathematics concept and, less frequently but still evident,

earning a grade they were not happy with (e.g. below a B) on an assessment. An example from

the data comes from Participant 13, when asked "What struggles, if any, have you experienced

with math?", she said: "Recently, the log stuff that we've been learning, it's been harder…but I figured it out by asking lots of questions…. [and] I just kept meeting with people in the library and kept going over things, which helped" (Interview 2). This quote illustrates her improvement in understanding logarithms by persisting through her struggle to reach success, something that the treatment course and program works to encourage. Additional examples from the Journals and Interviews can be found in Table 25.

Table 25: Examples of Participants' Statements for Success and/or Struggle

Data Source and Participant	Example	Explanation
Journal 1 Participant 19	"I just want to get better at math and be confidence [*sic*] while doing it. Almost every year I cry while doing it because I just don't understand it."	This statement illustrates how she reflects on her past experiences (struggle) that appear to have negatively impacted her.
Journal 4 Participant 18	"I tried my best and I did well. This taught me that I should [be] more confident in my math skills."	This statement appears to illustrate the improvement (success) he made from Journal 3 to Journal 4.
Journal 5 Participant 16	"I feel better about math because at least now I <u>understand</u> what we're doing instead of feeling completely lost."	This statement appears to illustrate how she improved (success) across the semester.
Interview 1 Participant 8	"When [math] starts getting letters put into it and you know, my God. Fractions, not very good at….And word problems are the death of me."	This statement illustrates how he appears to negatively reflect on his past experiences (struggle).
Interview 2 Participant 10	"When we're doing all the logarithms….it felt really easy to me and it came easy and it proved it on the test that I knew what I was doing and I would help people because they would always come to me because I knew what to do."	This statement helps to illustrate the improvement (success) he appeared to make during the semester between Interviews 1 and 2.

These examples of success and/or or struggle, along with the frequency in which they were discussed by participants in Journals and Interviews, illustrate the great effect that these experiences can have on students.

125

It is important to highlight that the opportunities provided for students to feel successful and to persist through struggle is an element specific to the treatment mathematics course and the C4 Scholar Program. Inclusion of this element offers a different experience for students than if they had been enrolled in a traditional mathematics course at the university. This is because the course was intentionally designed to feel like a safe place for learning by building in supports to hopefully encourage student success and help them persist through struggles. These supports, introduced in Chapter 3, included a team of instructors who often meet to discuss student progress and issues, mathematics workshops that are run by the instructor instead of an undergraduate assistant, a study session built into students' schedules, and acceptance of revisions on student work when possible. It is important to note that these supports were rarely or generally reported by the students and will be discussed further in Chapter 5.

Course instructor. Second, the data from participants' self-reports indicates that the mathematics course instructor, who was chosen specifically for the program, could also have contributed to the observed changes in self-efficacy. A statement from the program coordinator illustrates that he was "interested in [choosing] someone with strong background in mathematics education, with particular interest in issues concerning affect" with developmental students, which is why the current instructor was chosen. According to the data, when students spoke or wrote about the instructor it was often paired with the student feeling welcome or comfortable enough to ask questions or seek help from the instructor and feeling supported by the instructor. No students spoke or wrote about the instructor in a negative way.

The data from the Interviews indicate that 17 of the 19 participants specifically mention support from the instructor and how it has affected them. For example, Participant 4 stated in Interview 2 that the instructor has "always been there for us [for] whatever kind of challenges.

126

She's always helped work with us rather than against us…I really enjoyed having her as a professor". Additionally, when asked if his feelings about math changed over time this semester, Participant 6 said, "Kind of. Like I said, I wasn't at first 'was like, math. Now it's, oh yeah, math, it's fun" (Interview 2). And, when asked what he believed contributed to those changes he said, "Lauren. Obviously, I like her as a teacher. The surroundings, my classmates and I think it's just them all brought together, math" (Interview 2). Another example of how the instructor can affect a student comes from Participant 15's response to the interviewer asking if his feelings about math has changed at all during the semester. He said, "No. Well, yes…Because I really don't like math, but Lauren made it really welcoming, like a welcoming feeling" (Interview 2). Each of these examples show how the course instructor can have a positive effect on students in relation to mathematics.

In another example, Participant 8 comments, "I've talked to Lauren about some of the struggles I've had and she's helped me through them" (Interview 2). This was in response to the question, "Did you ever talk with teachers, friends, parents, or other family about these successes or struggles?" The struggles to which Participant 8 refers were related to understanding his mistakes on exams once they were graded and returned. To help this student, I sat down with him and explained similar examples so that he could make sense of the concept. A second example comes from Participant 1, in his response to what he feels has contributed to his positive change in feelings toward mathematics over the semester:

> Just my confidence. I would say more of the professor, has been very inviting and if you need more help just come ask. Providing that extra support and that reassurance… like soothing me into thinking that I can go to her for problems or ask for help, and with her

not having like any problems with it….And it's just, it made me feel better about going to her and having more confidence in math. (Interview 2)

As the instructor, with evidence throughout the daily instructor journals, I tried to emphasize that students could come to me anytime to ask questions and that it was okay to ask questions throughout the semester. Participant 1's statement provides evidence that he recognized this sentiment. Further, both examples illustrate the support and positive effect that an instructor can have on individual students. For Participant 8, it seems that I helped him overcome struggles he was experiencing with the content in the course. For Participant 1, it seems that I provided support for him to be able to ask questions and gain confidence in mathematics. Both examples provide evidence of the instructor possibly having a positive effect on the self-efficacy of these students and, according to analysis, both students made gains in self-efficacy on the *MSES*, Interviews, and Journals.

Mention of the instructor also appeared in the Journals at least once from 13 out of 19 participants, even though there was no prompt that asked students to consider writing about the instructor. For example, Participant 7 wrote, "In this math course this semester I believe that I have improved a lot. I think this is because we had a teacher that was willing to help us and not get mad about it. I'm very thankful for that" (Journal 5). Additionally, Participant 13 wrote, "I like the way the professor teaches. She teaches slowly and allows time for questions to be asked. Because of that, it has made me feel more confident in my ability to do math and creates less self-doubt" (Journal 2). These quotes illustrate how participants see the instructor of a course and the importance they give to being able to ask questions and seek help from the course instructor. As stated previously, no participant spoke or wrote about the instructor in a negative way.

Peer support. The third most prominent topic is peer support in the learning community, which could have also contributed to the observed changes in self-efficacy. According to the data, students spoke or wrote about feeling supported by their peers and the importance of being able to go to their peers for help with coursework. In particular, 17 participants reported this topic, which is made possible by the learning community, an element that is intentional with the treatment mathematics course and C4 Scholar Program. The data collected from Interviews showed greater mention of this element than that of the Journals. This is, likely due, in part, to the nature of the Interview questions that asked about participants' observations of their classmates having success or struggle. For example, Participant 13 said that having conversations with classmates who she observed having success in math influenced her, "because they helped me see that there's potential to do good. If you ask the person who's doing the work correctly, they can help you figure out where you're getting stuck" (Interview 2). In addition, from these questions stemmed many discussions about support from peers due to the learning community. For example, Participant 12 said, "I feel like the C4 program is very helpful. So we kind of have a close bond where we can ask those questions and get help" (Interview 2). In this statement, she is indicating an important component of the learning community which is the support of a community of peers and instructors. This is not something that is normally part of a standard mathematics course. Another example comes from Participant 7:

> I feel this program has showed me a lot, that it's okay to not be the best in the class and it's okay to ask questions about everything because it's just a learning process. As in high school you're like, "Oh well, if I ask questions I feel like I'm the dumb kid in class." But here, we're in a group of people that, we ask questions because we all don't understand it

129

at points. And, that there's people like you in the class where they're not judging you and you don't feel judged, so you feel like it's okay to be who you are. (Interview 2)

The effect of the learning community that the treatment mathematics course and C4 Scholar Program establishes is something that participants appear to see and report the benefits of in the data for the current study, which is reinforced by the statement from Participant 7.

Another example comes from Participant 4 who said, "I usually feel pretty good when I walk into class. I feel very positive because I know even if I do have a hard time, there's going to be others around me that are willing to help" (Interview 2). It is clear that she sees the benefit of having the support of her peers and the learning community that has been established. A final example comes from Participant 6, who said:

My friends in the, the class, we talked about our struggles and we'd help each other get through. We had similar struggles…. They've helped me not to give up and just helped me think about things in a deeper way rather than, "Oh, it's just this one way." I'm open to learning different ways to figure out a solution. (Interview 2)

It appears, from this example, that he saw the benefits of the learning community in the form of having relationships with and support from his peers. As stated earlier, similar statements were made by the majority of the treatment group participants.

Remaining topics. There were three other elements discussed by students, but with less frequency. These included support from others outside the course (which is not specific to the course or program), general components of the C4 Scholar Program (which did not have a singular focus), and outside influences such as the class social drama (which is not specific to the course or program). These will now be briefly discussed.

The fourth most common topic was support from others outside the course. Six participants' Journals and 16 participants' Interviews indicated that they had support from family, friends, etc. and they often reached out to them for help or words of encouragement. This is not unique to the course or program because students outside the treatment course likely reach out to family, friends, etc. for support as well. The following example from Participant 4 illustrates how she reached out to her mother for support. In Interview 2, she said,

> When I didn't feel I did so good on my exam 3, I let my mom know, and she just told me, "It'll be okay, you'll do better next time…make sure you know what you're doing before you go and revise it."

The fifth topic involved some component of the C4 Scholar Program. Components that were discussed included being enrolled in the program in general, the study session, revisions, and talking with the program coordinator. For example, Participant 18 often took advantage of the study session that was built into the program. He said, "I go to the study hours a lot and that is a lot of help…. I recommend for more people to go to study hours" (Journal 3). This shows that this component of the program seemed to be supportive of his learning and may have positively contributed to his experience in the treatment course. It is important to note that mention of these components was seldom and often discussed generally by participants. The sixth most common topic was related to some outside factor that affected the student. Only one participant made a comment in the Journals and only eight participants mentioned a factor of this type in the Interviews. This is not unique to the course or program because students outside the course likely have outside factors that affect their experience throughout the semester as well. For example, Participant 2 said, "Math has been a little hard lately and the last two weeks have been quite rough for me, just due to some problems in class and outside problems" (Interview 2). This

shows that the outside factor, he was likely referring to the class drama in which he was directly involved, may have affected his work in the treatment mathematics course. Other examples of outside factors included being involved in a court case, having a job, being involved in an extracurricular group on campus, and pressure from friends to go out instead of doing homework.

In summary, changes in self-efficacy for mathematics were evident based on the analysis of participants' self-reports in Journals and Interviews. The discussion of the participants and presentation of results shows that everyone had a unique experience in the course. Analysis indicates that eleven participants showed gains in self-efficacy on both the Journals and Interviews from the beginning to end of semester, with some fluctuations throughout. The average gain shown by participant Journals is 1.58 levels and the average gain shown by participant Interviews is 1.16 levels. Additionally, no students showed no change on both measures and no students showed a decrease on both measures. Two participants showed a combination of increase on one measure and decrease on the other measure, three participants showed a combination of increase on one measure and no change on the other measure, and three participants showed a combination of decrease on one measure and no change on the other measure. Further, the number of variable statements made by participants, on average, increases throughout the semester and shows that students may have been more aware of changes in their experience as the semester progressed. Additionally, the most commonly discussed topics that were reported with these observed changes, as discussed by the participants, included success and/or struggle with mathematics, the course instructor, and peer support in the learning community. Other topics, but on a smaller scale, included support from people outside the

course, general components of the C4 Scholar Program, and outside influences such as the class social drama.

Presentation of Course Assessments and Instructor Perspective Results

Prior to discussing the results of the fourth research question, it is important to present the results of students' earned grades and the Instructor Perspective because they have not been previously discussed like the other data sources.

Course assessment results. Table 26 contains the results of the four exams (Exam 1, Exam 2, Exam 3, and Final Exam) for each of the 19 treatment participants. These are the initial earned scores. Please note that red exam scores indicate a significant drop (10 or more points from previous score) or a score below 70%. If "Revised" appears next to a participant's scores, that means that the student revised at least one of those assessments. A revision takes the average of the initial score and the revised score to determine a final score. For example, if the student earned 80% and then earned a revised score of 100%, the student's final score would be recorded as 90%. Revisions, as discussed in Chapter 3, have always played a role in the C4 Scholar Program.

Table 26: Results of Treatment Participants' Earned Exam Scores

Participant #	Exam Scores (out of 100)	
1	97, 99, 93, 95	Revised
2	99, 68, 74, 58	Revised
3	78, 58, 91, 86	Revised
4	96.5, 96, 70, 93	Revised
5	100, 96, 100, 96	Revised
6	96, 81, 84, 95	Revised
7	87, 79, 85, 92	-
8	91, 91, 82, 96	Revised
9	97, 79, 81, 91	Revised
10	94.5, 94.5, 95, 95	Revised
11	96, 98, 99, 95	Revised
12	97, 94, 95, 93	Revised
13	100, 95, 87, 93	Revised
14	85, 81, 55, 62	Revised
15	77, 73, 32, 66	Revised

Table 26 (cont'd)

16	95, 80, 55, 87	Revised
17	96, 63, 66, 73	Revised
18	82, 69, 86, 89	-
19	93, 91, 89, 85	Revised

Additionally, Figure 13 shows a visual representation of the treatment group's exam scores.

Even though scores are independent of one another, the lines between scores aid the reader in

tracking a student's performance from Exam 1 to the Final Exam. Moreover, the lines help to

visually represent gains, losses, or consistencies in scores for each participant.

Figure 13: Exam Scores for Treatment Participants and Trends Between Scores
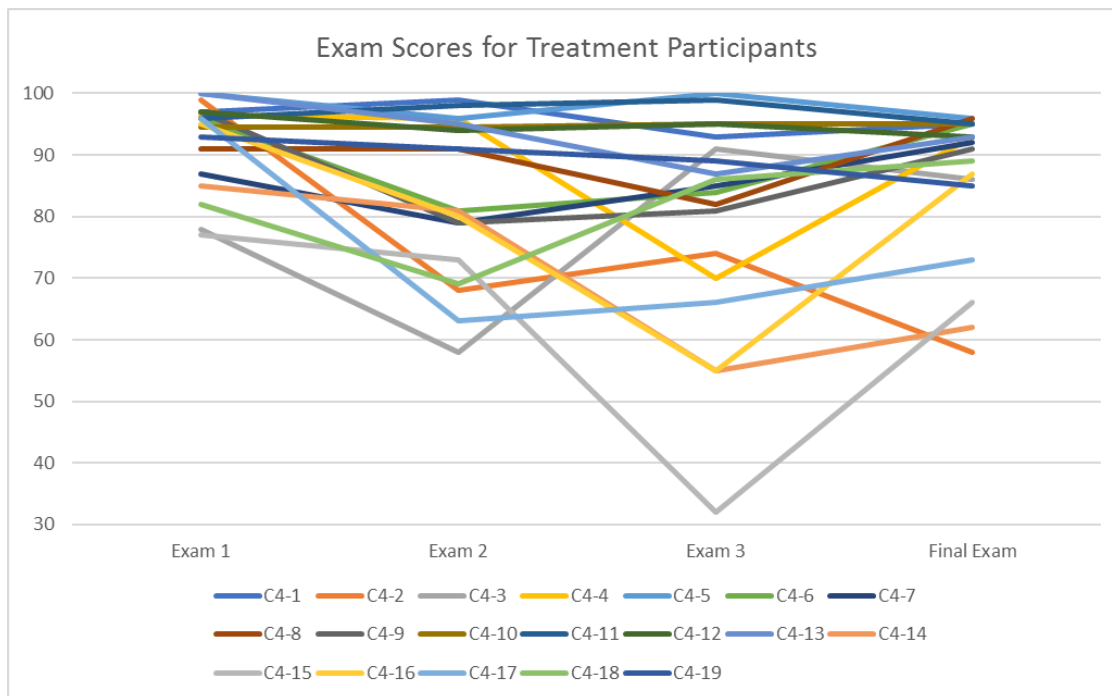
Instructor perspective results. Recall from the data analysis description in Chapter 3,

the Instructor Perspective results were formed by the analysis of documented instances of

possible influence on participants' self-efficacy throughout the semester in daily instructor

journals. The number of positive and negative entries about an individual in the instructor

journals were considered when assigning a rating for each period of time for each participant.

The rating scale is provided again in Figure 14. The results of this analysis are presented in Table 27 and the number of positive and negative statements in the form of (positive, negative) is included below the rating in each cell. It is important to note that if students had two or fewer entries for a specific time period, their cell in the table did not receive a color but they were still appropriately labeled based on the available data. These occurrences should not be surprising because instructor journaling can be a difficult task, as it was acknowledged in Chapter 3.

Figure 14: Self-Efficacy Scale for Instructor Perspective
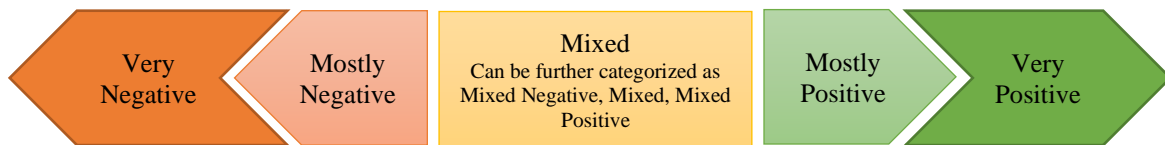
Table 27: Results of Instructor Perspective for Treatment Participants

Participant	Instructor Perspective (Week 5, Week 10, Week 16)		
1	Mostly Positive (5, 2)	Mostly Positive (9, 3)	Very Positive (13, 2)
2	Very Positive (2, 0)	Mixed – Positive (3, 2)	Mixed-Negative (4, 7)
3	Very Negative (0, 3)	Mostly Negative (2, 5)	Mostly Positive (7, 2)
4	Mostly Positive (4, 2)	Mostly Positive (5, 2)	Mixed (4, 3)
5	Mixed-Positive (5, 3)	Mixed-Positive (3, 2)	Very Positive (15, 0)
6	Mixed (0, 0)	Very Positive (3, 0)	Very Positive (6, 1)
7	Mostly Positive (2, 1)	Mixed (3, 4)	Mixed-Negative (3, 5)
8	Mixed (3, 3)	Mixed-Positive (6, 4)	Mixed-Negative (5, 8)
9	Mixed (1, 1)	Very Positive (1, 0)	Mostly Positive (5, 1)
10	Mostly Positive (2, 1)	Mostly Positive (4, 2)	Very Positive (11, 2)
11	Mostly Positive (3, 1)	Mostly Positive (4, 1)	Mostly Positive (10, 2)
12	Mixed (1, 1)	Mixed (2, 2)	Mixed-Positive (5, 3)
13	Mixed (2, 2)	Mixed-Positive (3, 2)	Mixed (5, 4)

135

Table 27 (cont'd)

14	Very Negative (0, 1)	Mixed-Positive (3, 2)	Very Negative (0, 3)
15	Mostly Positive (4, 2)	Mixed-Negative (4, 6)	Mostly Negative (3, 7)
16	Mixed (1, 1)	Mixed-Negative (2, 3)	Mixed-Negative (4, 6)
17	Mixed (2, 2)	Mostly Negative (2, 8)	Mostly Negative (2, 9)
18	Mostly Negative (1, 5)	Mixed (9, 7)	Mostly Positive (9, 4)
19	Mixed (3, 3)	Mixed-Positive (3, 2)	Mixed-Positive (3, 2)

Research Question 4: Comparing the Data Sources

Research Question 4 is: "What consistencies and/or or inconsistencies exist between the results of each of the collected data sources: (a) the standard instrument (*MSES*), the students' perspectives through (b) journals, (c) interviews, (d) students' course assessments, and (e) the instructor's perspective, which was motivated by documented instances of possible influence on participants' self-efficacy throughout the semester in daily instructor journals?" To answer this question, ten comparisons were made in order to assess the level of agreement between each pair of data sources. The ten pairings are listed below, and the results will be discussed in this order.

1. *MSES* and Interviews

2. *MSES* and Journals

3. *MSES* and Course Assessments

4. *MSES* and Instructor Perspective

5. Interviews and Journals

6. Interviews and Course Assessments

7. Interviews and Instructor Perspective

8. Journals and Course Assessments

9. Journals and Instructor Perspective

10. Course Assessments and Instructor Perspective

As a reminder from the data analysis description in Chapter 3, "Strongly Agree" indicates an exact or close to exact match between the trends of the two data sources being compared. "Somewhat Agree" indicates a similar overall trend between the two data sources being compared, but with some fluctuations within the data that do not support a strong agreement. "Somewhat Disagree" indicates a weak, opposite overall trend between the two data sources being compared. "Strongly Disagree" indicates an exact opposite overall trend or an inverse relationship between the trends of the two data sources being compared. I will now discuss the ten comparisons, beginning with agreements and then discussing the participants for whom disagreements resulted. The tables that explicitly show all the comparisons are in Appendix K.

Agreements. Of the total 190 comparisons that were made between each of the data sets for the 19 participants, 160 comparisons showed agreement whether rated "Somewhat Agree" or "Strongly Agree". Given this report from data analysis, it was determined that the strongest agreement resulted between two pairings: Course Assessments and Journals AND Course Assessments and Instructor Perspective. This was based on the highest number of agreements between participants' data sources. In addition, the least agreement resulted between three other pairings: *MSES* and Journals, *MSES* and Instructor Perspective, and *MSES* and Course Assessments. Here is the ranking in order from most agreement to least agreement, with a count of number of participants whose data sources agreed out of 19 participants in the treatment group:

1. Course Assessments with Journals (19)

    Course Assessments with Instructor Perspective (19)

2. Journals with Instructor Perspective (18)

137

Course Assessments with Interviews (18)

3. Interviews with Instructor Perspective (17)

4. Interviews with Journals (15)

   MSES with Interviews (15)

5. *MSES* with Journals (13)

   MSES with Instructor Perspective (13)

   MSES with Course Assessments (13)

This shows that, for the most part, results of the data sources agree when reporting perceived

levels of self-efficacy for participants in the current study. Some comparisons show more

consistent agreement than others, and the disagreements that resulted will be discussed in the

next section.

   Disagreements. It is interesting that there are seven students for which all disagreements

between data sources occurred. These are Participants 2, 4, 5, 7, 15, 16, and 17. These

participants had actions in class, attitudes, statements in Journals and/or Interviews, work ethic,

motivation, and/or assessment scores that were observed to not coincide with some or all of the

other data sources. Each of these participants will be discussed further to detail the

disagreements that resulted and to shed light on possibly why these disagreements might exist. In

addition, to offer a further examination of these disagreements, I compared which data sources

agreed for these students and which disagreed to see if patterns emerged. Please note that for the

12 participants whose data sources all agree, Figure 15 illustrates the relationship between the

five data sources. A line connecting two data sources means that agreement of results exists,

while no line means that disagreement exists. Ten comparisons were made, and so ten possible

agreements (or disagreements) could exist.

138

Figure 15: Relationship of Data Sources for Participants Whose Results Agree
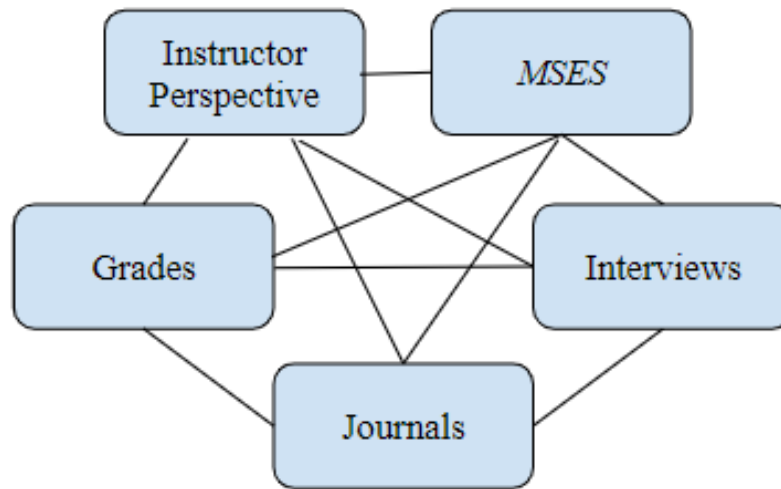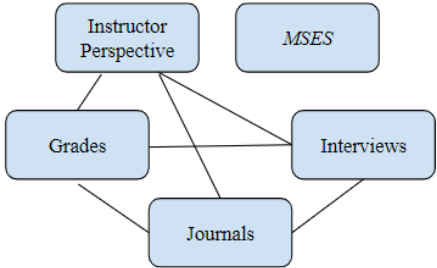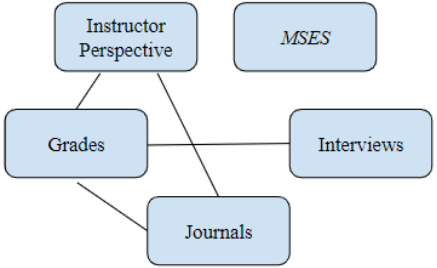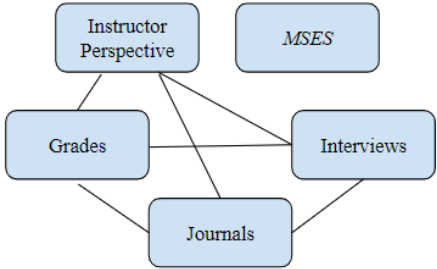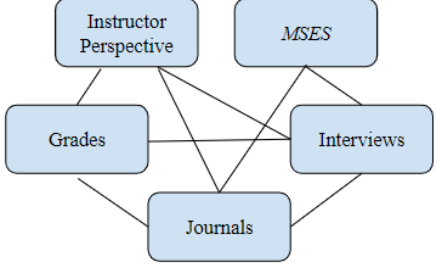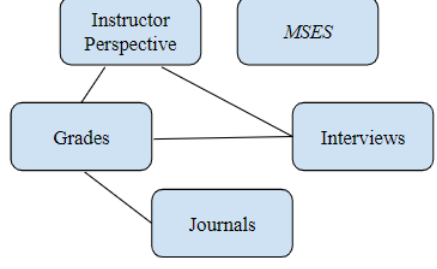
Table 28 explicitly shows the agreements and disagreements for each participant, which are
shown in their individual diagrams.

Table 28: Participants for Whom Disagreements Exist

Participant	Diagram to Illustrate Assessment of Agreement	Unique Factors
2		This student was engaged in many social distractions starting around Week 7.
4		This student was dealing with court, in favor of her safety, toward the end of the semester.

Table 28 (cont'd)

5		This student's first language is not English.
7		This student was diagnosed with an illness mid-semester that took her out of the classroom for almost a week, but the illness persisted for a few weeks after.
15		This student scored the lowest recorded exam grade of the semester (Exam 3 = 32%).
16		This student has endured many past family issues and was admitted for brief psychiatric care toward the end of the semester.
17		This student has dyslexia.

To summarize Table 28, the standard instrument *MSES* is the data source that has the most frequent number of disagreements with the other sources, contributing to 22 out of the total 30 disagreements. Of the remaining eight disagreements, seven of them involved the Interviews with another data source and the final disagreement is between Journals and Instructor Perspective. For Participants 2, 5, and 15, all disagreements between the data sources point toward *MSES* as the outlier. In addition, the Interviews were also an outlier for only Participant 2. For Participants 4, 7, 16, and 17, the disagreements did not all coincide to produce a clear outlier(s). For Participant 4, the varied relationships of Journals, Interviews, and *MSES* are the cause of her data sources not producing an outlier(s). For Participant 7, the *MSES* is an outlier but the match between Course Assessments and Interviews is causing an issue. For Participant 16, the relationship between *MSES*, Course Assessments, and Instructor Perspective is the issue for not producing an outlier(s). For Participant 17, the *MSES* is an outlier but the match between Course Assessments and Journals is causing an issue. Each of the participants who showed disagreements between data sources endured factors that were unique to themselves as individuals, which are mentioned in Table 28. This provides further evidence for how each participant's experience in the treatment course is unique. To further illustrate the unique experience of each participant, I will now discuss the seven participants for whom disagreements resulted. They will be presented in order of participant number.

Let us begin with Participant 2. According to the analysis results of his Journals, Interviews, and *MSES* scores, this participant began the semester with a level of self-efficacy that was much higher than any of his classmates, almost uncharacteristic of a student in an at-risk course. In addition, the Instructor Perspective supports that he began with the highest level of self-efficacy of any student even though fewer than three entries were made during daily

instructor journaling for this participant for the first rating. In particular, his *MSES* 1 score was higher than those of his classmates and he was the only participant to begin the semester with a Very Positive rating on both the Journals and Interviews (none of the other participants began with Very Positive on either). Yet, according to the analysis results of his Journals, he ended the semester with the lowest level of self-efficacy rating of all participants in the treatment group. His course assessments show decline in performance (and mastery) throughout the semester. On Exam 1 he scored 99%, on Exam 2 he scored 68%, on Exam 3 he scored 74%, and on the Final Exam he scored 58%. Further, according to the daily instructor journals, this student seemed (a) to lack motivation in class as the semester progressed because at the beginning he was observed to volunteer an answer for nearly every question asked to the group and was seen helping classmates with their work during class, (b) was frequently finishing assignments at the last minute as the semester progressed which was observed when the instructor would walk around the room collecting homework at the beginning of class, and (c) he was frequently engaged in social distractions from around Week 7 until the end of the semester. The ratings from the Instructor Perspective suggested a shift from Very Positive to Mixed-Positive and ending at Mixed-Negative. His Journals, which agreed with the Instructor Perspective and course assessments, indicate a decline from Very Positive to Moderately Negative across the semester. The analysis of his *MSES* and Interviews show that he remained Very Positive or in a very positive related category throughout the whole semester. This is a stark contrast from the results of the other three data sources. What is unique to this individual is his level of involvement in the cohort's significant and ongoing social drama of the semester, detailed in Appendix D. Other students were involved, but not near to the extent that this student was, which may be a factor in why these disagreements exist. A possible byproduct of his involvement was the appearance of

142

decline in his dedication, performance, and ultimately his self-efficacy in the treatment course, which appeared to be evident in his Journals, his grades, and the Instructor Perspective. It is interesting that around the time that he became heavily involved with the social distractions is when differences in the results of the five data sources became apparent. Around Week 7 is when the first documented social distraction occurred. This timeline correlates with the first reported differences in his results which was a decline on Journals, Instructor Perspective, and his Exam 2 score, while the *MSES* and Interviews remained very high from beginning to end. One might expect a decline to occur for all sources, but the decline that occurred was for data sources that focus on specific and smaller units of time, whereas the *MSES* and Interviews ask the participant to reflect on a larger scale. Perhaps the time period to reflect on was the defining factor that accounted for differences in results for Participant 2. The struggles that this participant faced, which were possible consequences from the social distractions, might have been more apparent in smaller time frames and less apparent in larger time frames.

Participant 4, according to daily instructor journaling, seemed to be consistently confident in mathematics from the start of the semester but then appeared to decline slightly toward the end. This is evidenced by her apparent good work ethic at the beginning of the semester paired with statements indicative of positive self-efficacy. For example, she appeared to be extremely focused during class while working, she would actively participate in groupwork, and she would stay after class to ask questions. An example of a statement as recorded in the daily instructor journaling was, "Oh! I know how to do this!" while she worked on an in-class activity and she said this with excitement. Her actions transitioned to more absences from class and apparent struggle, which was occasionally expressed by her and documented in the daily instructor journals, toward the end of the semester. For example, when turning in the last quiz of

the semester, she told the instructor she did not think she did very well. These observations informed the Instructor Perspective, which shows that she spent most of the semester at a Mostly Positive rating but then declined to Mixed at the end of the semester. Considering the results of the other data sources, her Journals show a shift from Somewhat Positive up to Moderately Positive at the middle of the semester. Then, the student declined to Neutral to finish the semester off. What appears in her Journals is a stark (and complete opposite) contrast from the *MSES* which shows that the student began at a "neutral" related category of Some Confidence and increased greatly to end at Complete Confidence. Also disagreeing with the results of her Journals, her Interviews show that she began at Neutral and ended at Somewhat Positive. What is unique to this participant is her involvement in outside legal issues that dealt with her safety toward the end of the semester. This appeared to be quite distracting for her because of her absences and developing struggles in the course. The analysis of the results from her data sources indicate that there is no clear outlier, but the issue lies with lack of agreement between her Journals and both the *MSES* and Interviews. Other than these two disagreements, agreement existed between all other comparisons. It does not appear that the legal issues, which were unique to her as a participant in this group, can offer a viable explanation for these two disagreements because if the legal issues affected her self-efficacy then one could expect a similar change to occur across all data sources.

Participant 5, according to daily instructor journaling, appeared to be (a) extremely motivated because of her preparedness in class and willingness to volunteer, (b) a hard worker because she would frequently stay after class to complete homework assignments immediately and/or ask questions, (c) consistently did well on her exams (earning 100%, 96%, 100%, and 96% on each of the four exams), and (d) appeared to have confidence that she would do well and

rarely appeared nervous. However, her *MSES* scores indicate that her level of self-efficacy was not overly high. In particular, her scores on the *MSES* show a very slight increase, but all within the same "neutral" related category of Some Confidence. The observations listed at the beginning of this paragraph, which informed the Instructor Perspective, ultimately agreed with the results depicted by all data sources except the *MSES*. Her Interviews show a shift from Neutral up to Very Positive. Her Journals similarly show that the student began at Neutral, then increased to Moderately Positive, and increased to Very Positive, which was maintained for the entire second half of the semester. What is unique to this participant is that English is her second language and she indicated struggle with this in the past to the instructor. This could explain why disagreements resulted only for her *MSES* scores because perhaps the instrument was written in such a way that her understanding of the questions was unclear and her interactions with the instrument were minimal. The questions for the other data sources were written by the researcher, who the participant had many interactions with throughout the course and may have become accustomed to. In addition, perhaps the fact that some data source collection methods were coupled with in-person interactions (Instructor Perspective and Interviews) could have played a role for this participant as well.

Participant 7, according to the analysis of her Journal responses, began the semester with a Neutral rating and then experienced some highs and lows throughout the semester, ultimately ending with a Neutral rating and thus appearing to show no change. The analysis of her *MSES* and Interviews, however, tell a different story. Her *MSES* scores begin the semester with a neutral related rating, which she maintains until the end of the semester when she makes a drastic jump up to a very positive related category, having the second highest *MSES* 3 score in the group. Her Interviews show that she began with a Moderately Positive rating and then fell

slightly to a Somewhat Positive rating. What is unique to this participant's experience is that she

suffered from an illness during the middle of the semester which took her out of the classroom

for almost a full week and the illness persisted for weeks after. No other participant was absent

this number of days in a row for any reason. Looking closer at her Journals shows evidence of

the perceived effects of this illness, although she never explicitly mentions the illness itself. She

experienced some success early in the treatment course, which is likely why the analysis of her

Journals shows an early increase to Moderately Positive but around the time that Journal 3 was

administered, with effects of the illness lingering for some time after, her self-efficacy appears to

drop throughout the semester to a Somewhat Negative rating. She is, however, able to recover

and the analysis shows she ends with a Neutral rating. The perceived effects of this illness are

also apparent in the Instructor Perspective, for which the analysis shows that her self-efficacy

appears to drop throughout the semester. Further, her grades on exams begin at a B+, then drop

to a C+ around the time of her illness, which also could have affected her self-efficacy. The

results of her Journals, course assessments, and Instructor Perspective seem to agree. The

Interviews and *MSES* appear to be outliers, although there is an agreement between the analysis

of her Interviews and course assessments, where her exam scores support the slight decrease in

self-efficacy that the Interviews show. One might expect similar changes in self-efficacy to occur

for all sources, but the data sources that generally agreed were those that focus on specific and

smaller units of time, whereas the *MSES* and Interviews ask the participant to reflect on a larger

scale. Perhaps the time period to reflect on was the defining factor that accounted for differences

in results for Participant 7. The struggles that this participant faced, which were possible

consequences from the illness, might have been more apparent in smaller time frames and less

apparent in larger time frames.

Participant 15, according to daily instructor journaling, appeared to start the semester off with a strong motivation to succeed. This was evidenced from him asking a lot of questions, staying after class, asking for extra practice problems, and seeking help when needed. His initial results from the data sources support his apparent confidence at the start because these recorded observations informed the first Instructor Perspective rating of Mostly Positive and the *MSES* showed an initial level of Much Confidence (being the third highest in the group). But after the second exam and especially in unit 3, he appeared as though he wanted to give up during class because of his apparent struggle with material we were covering. His course assessments, which were 77%, 73%, 32%, and 66%, show drastic decline in scores (failing his last two exams), which does not support the results of the *MSES*. What is unique to this participant is that he earned the lowest exam score (32% on Exam 3) of any participant for any one of the exams. His Journals show that the student began at Neutral, then decreased to Moderately Negative, then increased sharply to Moderately Positive, and then started a downward trend for the remainder of the semester until the student reached Somewhat Negative. Also, his Interviews show that he remained at Neutral all semester. The analysis of results for Participant 15 indicates that *MSES* is the outlier. The *MSES* appears to show that his self-efficacy was fairly high and even increased slightly over the semester, all within the category of Much Confidence, which does not agree with the results of any other data source for this participant. Perhaps this participant's strong desire to succeed at the beginning of the semester was carried through on all administrations of the *MSES* because this data source did not ask for direct reflection or explanations for his responses. The Journals and Interviews asked him to reflect and explain his responses, the Instructor Perspective was a collection of observations from the instructor/researcher, and the course assessments revealed performance; each of which either request or show direct evidence

of events that occurred and/or impacted the participant to shape his changing self-efficacy.

Additionally, the *MSES* was the only instrument that remained the exact same in form

throughout the semester, which may have helped to preserve similar responses from this

participant. Other data sources asked questions that might have been similar or related, but not

exact in each administration as was the case for the *MSES*.

Participant 16 appeared to gain self-efficacy early in the semester and then appeared to

lose this momentum and possibly lost some self-efficacy in the second half of the semester,

based on the entries from daily instructor journaling. The analysis of results for this participant

does not point toward a clear outlier. Overall, there is agreement between her data sources, but

there is lack of agreement between *MSES*, Instructor Perspective, and course assessments, with

the *MSES* being involved in both disagreements. In particular, this student started off very well

with her Exam 1 score, but her second and third exam scores decreased by 15 points and then

again by 25 points, respectively. This decline in scores does not seem to be reflected in her

MSES results. The *MSES* shows that the student started in a "neutral" related category of Some

Confidence and then increased and ended with Much Confidence for the second half of the

semester. The results of the Instructor Perspective, which agrees with the course assessments but

not the *MSES*, shows that this participant remained Mixed throughout the semester, even dipping

down to Mixed-Negative for much of the end of the semester. What is unique to this student is

her transparency in her discussion about having a lot of family-related issues that affected her

outside the classroom, and she revealed to the instructor that she was placed in psychiatric care

for a short period of time toward the end of the semester. It does not appear that her unique

circumstances can offer a feasible explanation for these two disagreements because if these

issues affected her self-efficacy then one could expect a similar change to occur across all data sources.

Participant 17, according to daily instructor journaling, fell asleep in class on occasion, finished assigned work at the last minute, attempted to get instructor assistance on nearly every problem assigned, had varying exam scores that were not overly indicative of mastery, and appeared extremely nervous in testing situations (in particular, he had sweat dripping from his forehead during Exam 3). These observations, which informed the Instructor Perspective ratings of Mixed and falling to Mostly Negative for the rest of the semester, are consistent with the analysis of the results of his course assessments and Interviews. For example, his course assessments show that he failed two out four exams (i.e. he earned 96%, 63%, 66%, and 73%) and his Interviews show that the student began at Moderately Positive and declined to end the semester at Neutral. His *MSES* scores, however, indicate that he had a high level of self-efficacy throughout the entire semester. In particular, he had the second highest score of his classmates on the first administration with a level of Much Confidence, he was tied for the third highest score on the second administration with a level of Much Confidence, and he had the fourth highest score on the last administration with a level of Much Confidence. His Journals interestingly show that he started at Neutral, increased to Somewhat Positive and then Very Positive, then significantly declined to Very Negative, and ended up increasing to Somewhat Positive at the end. This shows that the *MSES* and Journals, for the most part, appear to be outlier data sources for this participant. It is interesting that his Journals results only seem to agree with his course assessments and no other data sources. What is unique to this student is that he has dyslexia, which was revealed to the C4 program coordinator and instructors. The two data sources that were considered outliers for this participant, *MSES* and Journals, were written measures. Perhaps

this, paired with his dyslexia, is the cause for disagreements between some of his data sources. The Interviews asked for verbal responses and the Instructor Perspective was comprised of observations. The only other data source that was written were the course assessments, which do reflect struggle throughout the semester.

Discussion of these results reveal a pattern: for participants who have very few disagreements and do not show a clear outlier, their unique circumstances do not appear to offer an explanation for any differences in results. In particular, this applies to Participants 4 and 16, who both had only two disagreements. Additionally, a related pattern results for participants who have four or more disagreements and for whom clear outliers can be identified. In particular, a possible argument about the existence of disagreements could be made for Participants 2, 5, 7, 15, and 17. For Participants 5, 15, and 17, their unique circumstances could help explain their disagreements which generally centered around the *MSES* (and Journals for Participant 17 only).

For Participants 2 and 7, their results point toward differences in focused versus broad time-period reflections and suggest that when participants are asked to think about their experiences with mathematics for a longer period of time (one that may be undefined, i.e. before you came to college), details may be left out and so they may not be considering or reporting all of the events and elements that have shaped their current belief. Thus, important details about an individual's journey may be lost without the inclusion of measures that are collected more frequently. Measures that ask for reflection from participants can prove to be difficult for anyone because there are different major events that stick out in individuals' minds that could have varying effects on those reporting or reflecting on them. There may also be distracting and seemingly unrelated events that can impact an individual's responses. In particular, Participant 2's more frequent data sources (Journals, course assessments, and Instructor Perspective) all

150

agree and illustrate his unique experience. The Interviews and *MSES*, which were only collected twice, did not tell the same story, and appeared to lose details about his journey. In addition, Participant 7's Journals provide details of the journey that this participant endured throughout the semester because she is able to reflect on smaller units of time, rather than only thinking about herself and mathematics on a larger scale. Thus, important details about her experience during the semester were preserved which may have otherwise been lost.

For the twelve participants for whom analysis did not reveal disagreements, inclusion of the more frequently collected Journals still provides details about each individual's journey that may not have emerged in other collected measures. Further, the Journals may provide evidence of fluctuation in self-efficacy that illustrates the journey that each participant took throughout the semester that may not have been represented by the results of the Interviews or *MSES*. For example, Participant 9's Journals show fluctuation that the Interviews and *MSES* do not. Even though the data sources show an overall increase in self-efficacy for this participant, if data was only collected on her Interviews and *MSES* scores then the details of her initial gain, followed by slight loss, and then final gain in self-efficacy throughout the semester would not have been available. Therefore, the inclusion of the Journals helped to provide more details about her experience and a deeper understanding about this participant's journey results.

In summary, the *MSES* and the Interviews asked participants to think about themselves and mathematics more holistically while the Journals asked participants to think about themselves and mathematics in smaller timeframes. The course assessments and Instructor Perspective, although not reported by the students, are also indicative of smaller timeframes and more often agreed with Journal responses. Therefore, because of the frequency in which they were administered, the inclusion of Journal responses as a data source provides specific details

regarding participants' self-efficacy changing throughout the semester, that may not have emerged in the other data sources. The knowledge of these additional details enriches the understanding that can be gained about each participant's self-efficacy and how or why it may have changed throughout the semester.

CHAPTER 5: DISCUSSION

This chapter begins by summarizing the key findings for each of the four research questions. These findings will then be discussed in terms of the factors that likely shaped them and how those findings relate to the in current research literature on efforts to increase self-efficacy at the college level. To conclude, implications and limitations that stem from the current study will be identified and discussed.

Overview of Key Findings

The results of the first research question indicate that all participants in the treatment group made gains in their total self-efficacy scores on the *MSES* from the first administration of the survey to the last administration upon completion of the treatment course. Further, the Everyday Math Tasks subscale had higher results than the Math Courses subscale for the treatment group. The results of the subscales help to illustrate, in more specific ways than the total scores, how the self-efficacy of participants changed after the treatment. Coincidentally, the material covered in the course aligned fairly well with the questions in the Everyday Math Tasks subscale because of the common focus on the mathematics that one applies to daily life, such as financial math. The course was designed prior to selecting the survey and was taught as it was the previous year, so it was not designed based on the survey.

The results of the second research question indicate that upon completion of the treatment course, statistically significant differences resulted between the initially comparable two groups based on the improvements made across time on the *MSES*, all in favor of the treatment group, for total mean scores and the two average subscale scores, as shown in Table *29*. It was determined that the treatment group made greater gains at the end of the semester than those made by the control group, because the time-by-treatment interaction was significant for the total

MSES scores and both subscale *MSES* scores. These results suggests that the treatment course had a significant effect on the self-efficacy of the treatment group, because even though both groups started at similar points for their total *MSES* and two subscale scores, the treatment group showed greater improvements than those made by the control group at the end of the semester.

Table 29: Comparison of Total and Subscale Scores

	Treatment		Control	
	Beginning	End	Beginning	End
MSES Total	4.79	6.86	4.81	5.66
MSES Part 1 Subscale	4.93	7.46	5.41	6.58
MSES Part 2 Subscale	4.64	6.18	4.15	4.63

The results of the third research question show that according to analysis of Journals and Interviews, gains in self-efficacy were made from the beginning to end of the semester by eleven of nineteen treatment group students with some fluctuations throughout. Additionally, no students' self-efficacy remained constant on both measures and no students' self-efficacy decreased on both measures. Two students' self-efficacy increased on one measure and decreased on the other, three students' self-efficacy had no change on one measure and decreased on the other, and three students' self-efficacy had no change on one measure and increased on the other. The increase in the number of variable statements over time suggests participants were more aware of changes in their experience as the semester progressed. In addition, these data sources point to three intentional elements of the C4 Scholar Program and the treatment course that students reported as influencing how they felt about mathematics: (a) opportunities for success and the help to persist through struggle, (b) a course instructor who invited questions and was willing to help, and (c) peer support in the learning community. There were three other elements (support from people outside the course, general components of the C4

Scholar Program, and outside influences such as social issues) discussed by students, but with less frequency.

The results of the fourth research question show the number of agreements in resulting reports of self-efficacy from (a) the *MSES* survey, (b) student Journals, (c) student Interviews, (d) results of major course assessments, and (e) the instructor's perspective. The ten pairings of data sets that resulted from these five sources were compared for each of the 19 participants. The ranking in order from most agreement to least agreement (with the number of participants whose data was in agreement) is as follows:

1. Course Assessments with Journals (19)

   Course Assessments with Instructor Perspective (19)

2. Journals with Instructor Perspective (18)

   Course Assessments with Interviews (18)

3. Interviews with Instructor Perspective (17)

4. Interviews with Journals (15)

   MSES with Interviews (15)

5. *MSES* with Journals (13)

   MSES with Instructor Perspective (13)

   MSES with Course Assessments (13)

The treatment course seemed to be effective in increasing most participants' self-efficacy based on the analysis of data gathered from five sources. For the most part, the analysis of results of the five data sources agree, with some sources showing more agreement than others. Further, analysis of results suggests that for participants who have a small number of disagreements and do not show a clear outlier, their unique circumstances do not appear to explain any differences

in results. For participants who have four or more disagreements and for whom clear outliers can be identified, it seems that unique aspects of their personal circumstances suggest plausible explanations. It seems that the inclusion of Journals as a data source appeared to expose details about the participants' experience that may otherwise be lost. This finding emerged in the comparison of data sources and seemed to allow more knowledge and a deeper understanding to be gained about each participant and his/her self-efficacy. Based on these results, I offer the following discussion and implications.

Discussion of possible influences on findings. It is important to acknowledge that the results obtained in this study may be due to influences outside the treatment course. I will focus this brief discussion on recognizing influences that focus on the populations of the two groups in this study, factors outside of the treatment course, and influences that arose from the use of data collection instruments within the treatment course. First, it was mentioned in Chapter 4 along with the presentation of quantitative results that the populations of the two groups looked different, even though analysis showed there were no statistical differences. Specifically, it seems that the treatment group had noticeably higher participation from students who were male, White, and/or from rural high schools than those who participated in the control group, while students who were female, Black, and/or from urban high schools had higher rates of participation in the control group. Even though the results of the treatment could have varying effects on each individual participant, the treatment group already had students who were more inclined to have higher self-efficacy than those in the control group (i.e. higher participation from students who were White and/or male in the treatment). Direct support for this comes from Betz & Hackett's (1983) findings. It is important to recognize that perhaps the demographic characteristics of the two groups of participants was more supportive of the greater improvement

156

made by students in the treatment group even though demographic information is used as a baseline comparison (i.e. before treatment is applied).

Second, as was also mentioned in Chapter 4 alongside the discussion of effect size, it appears that something in the treatment had an impact on participants' self-efficacy. But it is also possible that factors outside of the treatment influenced these results. In particular, there could have been influences outside of my instructional decisions within the treatment course (i.e. more supportive nature of the C4 Scholar Program) that had an effect on the treatment participants, which is important to acknowledge. One example of these influences could include the encouragement and support received by a student from other instructors or the program coordinator, who are aware of student progress and/or issues because of the scheduled instructor meetings. Another example could include the C4 Scholar Program's focus on helping students move towards developing accountability and professionalism while enrolled.

Third, it is also important to recognize that the results obtained in this study may be due to influences from both the treatment course and the data collection measures (*MSES*, Interviews, and Journals). The data collection measures became part of the treatment course, making it difficult to separate them, and their use could have had an effect on the participants because they were being asked to actively reflect on their experiences throughout the semester. Interviews and Journals were part of the mathematics course and the C4 Scholar Program prior to this study, but the questions in those used for this study were altered to better align with the study's goals and could have influenced the results obtained from the treatment group. In particular, the data collection measures might better encourage students to recollect the successes and/or struggles they were experiencing within the treatment course. This could highlight and/or bring to the forefront the effects students experienced as a result (i.e. the Journals and Interviews ask students

to discuss successes/struggles and the *MSES* asks students about specific math tasks). It is unclear if the act of asking participants to reflect affects their self-efficacy, but the awareness that may result for participants could support positive or negative effects and might be reflected in their responses. Therefore, the presence of data sources might have encouraged students to reflect in more direct ways than they may have naturally done themselves which may have affected the results of their perceived self-efficacy.

Connecting Findings to Literature

As discussed in Chapter 4, three intentional elements of the treatment mathematics course and C4 Scholar Program were identified by students as influencing their feelings about mathematics and directly relate to Bandura's (1977) sources of self-efficacy. These are: (a) opportunities for success and the help to persist through struggle linking to mastery experiences, (b) the course instructor linking to verbal persuasion, and (c) peer support in the learning community linking to vicarious experiences. These elements appeared to offer more support for growth in students' self-efficacy in mathematics than what is available to similar students outside the program. It is important to note that treatment students' reports of these three elements were not direct references to self-efficacy, but they were most frequently discussed in students' Journals and Interviews and, ultimately, may be sources of self-efficacy for this group. This course was not designed around Bandura's (1977) theory on self-efficacy, but strong connections to the four sources of self-efficacy were present before data collection began and three of the sources emerged from the data.

The way in which students discussed these three elements strengthens the link to associated sources of self-efficacy. For mastery experiences, students primarily focused on their success or struggle with understanding mathematics content and it was discussed most frequently

158

by far. This result is not surprising because this is the most influential source and according to

Bandura (1977), "successes raise mastery expectations; repeated failures lower them" (p. 195).

This result was similarly found by Zientek et al (2019). The discussion from students that

focused on the support and encouragement they received from the instructor to ask questions and

seek help related to verbal persuasion. This ties directly to Bandura's definition of the source,

"People are led, through suggestion, into believing they can cope successfully with what has

overwhelmed them in the past," (1977, p. 198) because the encouragement students received

from the instructor to ask questions and ask for help, which also encouraged them to keep trying,

appeared to help them see and believe they could overcome their struggles. The data includes

some evidence consistent with the influence of verbal persuasion. In the Journals and Interviews,

the participants were not asked about the instructor but many of them discussed the positive

influence the instructor had on them. And, in the daily instructor journals, which informed the

Instructor Perspective, there are many recorded instances in which the instructor would

encourage individual students or the entire group. For example, if a participant expressed

confusion and/or frustration about a particular topic, the instructor's response would often

intentionally highlight what the student was doing well (i.e. they showed understanding of what a

fraction means) and encouragement to move forward. These conversations were always in

person, and rarely in email; this is important to distinguish because many methods are available

for communication, which is evident in the works reviewed for this study in Chapter 2. In person

conversations for the current study were likely due to the course design which included course

meetings four days per week and an additional workshop period. This is similar to the finding

from Olani et al. (2011), whose study observed increases in student self-efficacy, especially for

those who recognized support from their instructor. Finally, for vicarious experiences, students

expressed the importance of support from their peers and discussed how their interactions with peers influenced how they felt about themselves and mathematics. This links directly to the definition of this source which is, "inferences from social comparison" which came from observation and/or learning about their classmates' experiences in the course (Bandura, 1977, p. 197). This is a difficult source of self-efficacy to measure (Usher & Pajares, 2008), but the results in Chapter 4 provide support for the emergence of this finding. The arrangement of groups during class sessions helped to support the possible effect of vicarious experiences; and students' responses in Journals and Interviews shows evidence of this effect. For example, Participant 12 discussed how observations of classmates having success in math influenced her. She said, "I feel like some of my friends in this class have success. And I kind of pushed myself to be like those people in math because some parts they'll be doing really good and I struggle with" (Interview 2). This quote helps to illustrate how observations of one's peers can act as a positive influence. Groups, which were intentionally arranged by the instructor, paired some students who were doing well in the course with those who were not doing as well. The groups also paired students who had similar performance in the course. Thus, as was observed in this study, learning communities can have many positive effects on students (Love, 2012).

In Chapter 2, I discussed various studies that focused on implementing interventions for students' self-efficacy toward mathematics for pre-college and college student populations. I would now like to address how the current study fits in with existing research. First, this study focused on the college age population. Second, in terms of length of time, this study spanned one full college semester. Other studies included in the review with a focus on college students and discussed in Chapter 2, varied from one hour to one full college semester. Third, this study did not assess longitudinal effects of the intervention on self-efficacy. It only considered any

perceived changes made, from five data sources, from the beginning to end of one semester. Fourth, the current study could be classified under each of the four groupings that were used to synthesize existing intervention work: (a) focus on student study and motivation skills because of its involvement with the C4 Scholar Program, (b) focus on direct teacher influence because of my instructional decisions as the instructor and researcher, (c) focus on changes in course structure because the treatment course is different than what students would have received in the traditional course (as described in Chapter 3), and (d) focus on a change in curriculum (as described in Chapter 3). Fifth, it seems that this intervention was successful because most students in the treatment group appeared to show an increase in self-efficacy, especially when considerations of improvements were made in comparison to a comparable control group.

The results of the current study will now be discussed in relation to those of existing research reviewed in Chapter 2 on the basis of how the intervention affected participants, which of Bandura's (1977) sources were emphasized, and if the effects lasted beyond the intervention.

Effect of interventions. First, it appears that the current study demonstrated success in increasing most students' self-efficacy toward mathematics based on their participation in the treatment mathematics course. When compared to the results of existing research that was reviewed in Chapter 2, most of the studies reviewed in Chapter 2 similarly reported that an increase in self-efficacy toward mathematics for college students was observed (Cordero et al., 2010; Valencia-Vallejo et al., 2018; Brewer, 2009; Olani et al., 2011), although Brewer (2009) found that increases occurred for the control group as well. Further, Hekimoglu and Kittrell (2010) reported possible positive influence on self-efficacy, while Hodges (2008) and Hodges and Kim (2010) reported no effect at the conclusion of their studies. No studies reported negative effects.

161

Impact of Bandura's sources. The results of the current study indicate the emergence of three sources of self-efficacy (master experiences, verbal persuasion, and vicarious experiences) that appeared to be influential. Other studies mentioned or focused on the sources of self-efficacy, but not all had similar results. Of the four sources, verbal persuasion was most commonly studied by the researchers included in this review, followed by mastery experiences. Beginning with mastery experiences, the current study and Cordero et al. (2010), when coupled with self-persuasion which is not one of Bandura's sources, found this source to be influential. However, Brewer (2009) did not find mastery experiences to be influential. Next, the current study found verbal persuasion to be influential, which is similar to the findings of Valencia-Vallejo et al. (2018); although the current study found influence via in-person interactions while Valencia-Vallejo et al. (2018) completed their work online. This result contrasts with Brewer (2009), Hodges (2008), and Hodges and Kim (2010), who did not find their virtual work with verbal persuasion to be influential. The current study also found evidence for influence from vicarious experiences, but no evidence for emotional arousal. Similar findings, in part, were found by Hekimoglu and Kittrell (2010), who did not directly measure any of the four sources but were able to make connections to vicarious experiences and emotional arousal.

Stability of the effects of interventions. The current study did not consider the stability of the reported results. To contrast, two studies considered the stability of their results. Cordero et al. (2010) found that their positive results were present six weeks after the intervention, and Hodges (2008) found the intervention (email messages) to be influential initially, but the effects disappeared by the end of the semester after the conclusion of the four-week treatment. Therefore, the results of the current study seem to somewhat align with the results of existing literature. This conclusion is based on the comparisons of how the intervention affected

162

participants, which of Bandura's (1977) sources were influential, and if the effects lasted beyond the intervention.

Implications

I will now discuss implications that focus on implementing the treatment and/or program, measuring students' self-efficacy, and those for future research.

Implementing the treatment and/or program. Given that the treatment course seems to have had a positive effect on self-efficacy for the majority of students and intentional elements were highlighted by students in the data, institutions that seek to reduce attrition of at-risk students should pay attention to specific elements of the course and C4 Scholar Program that may have contributed to observed changes in self-efficacy. For instance, the data indicates that these elements (opportunities for success and/or struggle, the course instructor who is inviting of questions and who is willing to help students, and peer support in the learning community) cover three of the four sources of self-efficacy. This is important because these three elements are things that could help these students move towards experiencing and achieving success and, ultimately, improve their self-efficacy. This improvement in self-efficacy can increase their chance of persisting through college because, generally, students with a higher level of self-efficacy seem to persevere when faced with challenge and perform higher than those with lower levels of self-efficacy in mathematics (Gumml, 2013). Additional support for this finding comes from Lopez and Lent (1992), who recommend that developmental mathematics programs work to "provide students with genuine mastery experiences…[which] may enhance students' math self-efficacy" (p. 11). Ultimately, this could help students who are less likely to finish a degree program (Benken et al., 2015; Cafarella, 2016). These elements, which are explicitly supported in the literature, should be integrated into mathematics classrooms that host at-risk students

163

whenever possible and fuels the following implications for implementing the treatment course and/or program.

Institutions should encourage the integration of supports that provide more opportunities for success and help students persist through struggle in mathematics. These can take a variety of forms, but for the current treatment course and program the supports, which were first introduced in Chapter 3, included a team of instructors who often meet to discuss student progress and issues, mathematics workshops that are run by the instructor instead of an undergraduate assistant, a study session built into students' schedules, and acceptance of revisions on student work when possible. These supports, however, were rarely or generally mentioned by the participants and so the subsequent discussion notes how these may have affected the students. Even though minimal evidence emerged from the data, these were the supports that the C4 Scholar Program implemented in order to try to help students have opportunities for success and learn how to persist through struggle and so their intention is worth mentioning.

These supports may help to provide opportunities for success and may help students persist through struggle in the following ways: The instructor team meetings are used to help identify issues that individual students are having, how the instructors can best support the individual, and a plan to, hopefully, help the individual persevere. The mathematics workshop being led by the instructor instead of a student may help to further encourage students to see the instructor as a resource and provides the instructor with time, outside of normal class meetings, to strengthen skills that students may be struggling with, which may be more difficult for an undergraduate assistant to identify. The inclusion of a study session may help students learn to devote time to studying and working outside of normal class meetings. The acceptance of revisions on student work may help students to revisit past mistakes, which they may have

164

otherwise ignored, and work to correct and understand them. It is important to reiterate that these supports were rarely or generally mentioned by participants in the data, but the experiences of success and/or struggle that may have resulted from these were reported.

Further, the institution should be mindful when choosing instructors to teach this population. In the initial search for instructors of the C4 Scholar Program, the program coordinator indicated that those involved, "sought instructors with the ability to collaborate in an interdisciplinary setting, the potential to adapt their courses to align with program goals, and experience in working with developmental students." Based on the findings of the current study, institutions should also look for instructors who are caring and sensitive to the needs of at-risk students. This is supported by Higbee (1993) and by the data collected from the Journals and Interviews because of the importance students seemed to place on feeling comfortable enough to ask questions and seek help from their instructor. Further, the daily instructor journals, which formed the Instructor Perspective, provides evidence of many students feeling comfortable enough to approach the instructor to discuss personal, social, or academic related issues. Ultimately, the instructor can, of varying degree, help foster success or act as an obstacle to student success. Each of these components can help to foster a relationship between the instructor and student and is one that can be vital to the success of those who are considered at-risk.

Finally, peer support in the form of learning communities is also important and should be considered by institutions when working to increase students' self-efficacy towards mathematics. Evidence in support of this recommendation comes directly from the self-reports of students in their Journals and Interviews. The relationships that the majority of students formed with their peers appeared to be positive and supportive of their success in the course. Therefore, institutions

should strongly consider the implementation of learning communities for their at-risk students if they have not already done so.

There are costs, however, associated with the implementation of a program or components of a program such as the C4 Scholar Program. In addition to monetary costs used to support these students, instructors who agree to implement the program will also assume the added cost of time because of the workshops that are instructor led and regular meetings with the instructor team.

Measuring self-efficacy. The data collected from student Journals show the greatest level of agreement with other data sources, but it is important to recognize this does not mean it is the best measure of self-efficacy and agreement does not imply validity. However, the results of the current study discern that there appears to be a perceived benefit of including data collection methods that are more frequent (e.g. Journals) to determine and understand students' levels of self-efficacy for mathematics. This points to recommendations for instructors and researchers. First, for instructors, who are interested in measuring, understanding, and possibly increasing their students' self-efficacy toward mathematics, data collection or check-in methods that ask students to reflect over a small time frame (e.g. a single unit and/or once every four to five weeks) and not just at the beginning and end of a year or course should be considered. This strategy can give the instructor insight about his/her students by possibly providing more details about students' experiences that may not emerge from the sole use of less frequent measures. Further, this could support a deeper understanding about each participant's self-efficacy and how or why it may have changed throughout a year or course. The findings of the current study support using reflections of this type. For instance, the Journals were given five times across one 16-week semester and appeared to have the highest level of agreement with other data sources of

166

the three instruments completed by participants. Therefore, the frequency of reflective journals should be given at least once every four to five weeks during a unit in the instructor's course, as was done in the current study, but not too often (e.g. every week) to avoid participant fatigue. This practice might help instructors understand their students' current and changing levels of self-efficacy and may help to identify what students feel has impacted them. In addition, this could provide instructors with a more detailed picture of students' levels of self-efficacy throughout the semester and might bring instructors closer to knowing and understanding their students as unique individuals, as is shown in the results of this study. Collecting and reflecting on this data from current students might also help instructors learn to identify and respond to future students' self-efficacy in mathematics. It is, however, important to acknowledge that instructors have many responsibilities in the classroom and the recommendation of this additional focus may prove to be difficult because of time constraints, preparation of materials, and reflection on student responses to further understand this construct.

Second, researchers in the field should also consider check-ins with participants at least once every four to five weeks when collecting data to possibly provide a more detailed picture of students' self-efficacy toward mathematics. For the current study, analysis of the Journals, which was collected most frequently, seemed to offer a detailed level of understanding about an individual's current level of self-efficacy and what has impacted him/her. This could support researchers in learning more about how and why students' self-efficacy in mathematics may change over time and can also help to identify fluctuations that may not be measured by less frequent measures. Therefore, in addition to measures of self-efficacy that may be given only at the start and end of the semester or term, it is recommended that researchers collect data from participants at least once every four to five weeks as well so that important details do not become

lost. It is also advised to not check-in too often (e.g. every week) to avoid participant fatigue. As the researcher for the current study, I found that the results produced by the Journals appeared to provide great detail about the journey each participant took as they progressed through the treatment course. Further, this closer observation of self-efficacy can help researchers point toward practices to support students and instructors during positive or negative changes.

It is important, however, to state that even though it appears that the Journals offer more details about participants and their experiences, this study is not declaring that Journals are the best source of data on self-efficacy in mathematics. At times, this method of data collection may have more to offer than others, but at other times, Journals may not be the best choice. For instance, Journals might be a good choice for a study that is longer-term (i.e. roughly a semester or longer) or if the researcher is interested in how students themselves make sense of perceived changes. Journals might not be a good choice for a study that is shorter (i.e. less than a semester) or if the researcher is only interested in whether changes for students are occurring, in which case a survey might be better suited. It should be up to the researcher to decide what method is best for his/her work.

Future research. Future implications for research involve another rendition of the current study, but the perceived effects of the treatment course would only be measured with the *MSES*. It is important to note that the *MSES* is the only instrument that was not written by the researcher. Self-efficacy is not a physical characteristic and it can be difficult to measure and so, using a standard measure such as the *MSES* could help to produce results that are as objective in nature as possible. Doing this might provide insight that the current study does not because there would not be possible influence from the use of other data sources, which was identified previously in this chapter.

However, due to the unfortunate circumstances of the COVID-19 pandemic and the virtual nature of much coursework at the university as a result, the C4 Scholar Program did not enroll a cohort for the 2020-2021 academic year. If I were to move forward with this work when there is another cohort that enrolls at the university, I would like to incorporate what I learned from the current study. Based on the results of this study, I learned that students appreciate various aspects of the treatment course and the C4 Scholar Program including opportunities to experience success and/or persist through struggle, their instructor, and peers in the learning community. In order to understand more about how these components and possibly others specifically influence students, it would be interesting to ask future students to intentionally reflect on their experiences with these aspects of the C4 Scholar Program. This data, which could follow-up the current study, might provide insight that would not only benefit the mathematics course, but the C4 Scholar Program as a whole by providing more information from students about intentional elements.

In addition, another direction for future research is to establish more C4 Scholar Program cohorts at the university in which the current cohort is hosted and at other institutions around the country. This should be done in a specific order of steps. It would, first, be interesting to see if an additional cohort(s) at the current university would experience similar results of at-risk students' self-efficacy toward mathematics given a different team of instructors. This new team of instructors should be chosen given the criteria previously outlined by the program coordinator and also, based on the data of the current study, be open and encouraging to students asking questions or seeking help. Doing this would maintain a similar environment, one that has been established from the development of the program, while still interacting with the same C4 Scholar program coordinator. Second, if other institutions implemented a similar program it

169

would be interesting to see whether an institution's current (or lack of) efforts to support at-risk students has differing effects on the results of at-risk students' self-efficacy toward mathematics upon conclusion of their first-semester mathematics course. To do this, other institutions who have a history of trying to support at-risk students should first be involved. The reason for this is to try to mimic an environment similar to what has resulted from the nature and history of the C4 Scholar Program at the current university. This would help to shed light on whether this program can be extended to other locations where the C4 Scholar Program is not present, but where similar groundwork may have already been established. Once results have been analyzed and compared, it would then be appropriate for institutions who do not traditionally focus on the at-risk student population, to be involved and implement cohorts for their at-risk students. This step would help determine whether the program can be extended to any institution where there is none or little previous groundwork of support for at-risk students. If this step is found to be unsuccessful, then it might be appropriate for institutions without a past focus on at-risk students to establish a foundation first and then try to reimplement the program. Therefore, this implication for future research calls for the current study to be implemented again but in different stages to see if the C4 Scholar Program could be established at any institution, no matter the degree to which they have previously focused on the at-risk student population.

Limitations

The most important limitations of this study are (a) the relatively small number of participants, (b) differences between the two groups, (c) the possible impact of instructor bias and where it likely entered the data, and (d) the small numbers of statements to be coded in qualitative data at times. I chose to highlight these limitations because they can affect data collection (bias and participants) and data analysis (bias and coding) in major ways.

One important limitation was that the current study engaged a relatively small group of treatment participants. Future work involves a larger group and/or a whole class of participants, recall that four students in the treatment pool decided not to participate. These four participants chose not to participate in the current study for unknown reasons. I have mixed notions on whether I believe their participation would have changed the results because these students ended the course with varying degrees of performance and varying levels of perceived self-efficacy, based on daily instructor journaling. Two of the students who decided not to participate could be considered similar to those who participated because their levels of performance were similar to some participating students and their attendance was near perfect, meaning that all data items could have likely been collected. So, I do not believe that their participation would have altered the main results of the study. However, the other two non-participants, due to low engagement and/or poor attendance in the course, earned the two lowest final grades. It is likely that participation from these two students might have resulted in missing data and could have affected results in major ways. For instance, these two students might contribute more disagreements between data sources to the results because the students appeared, from daily instructor journaling, to be less engaged with the course. One had many missing assignments and many absences. The other struggled to stay focused in class and occasionally fell asleep during class. Because of their engagement issues, I am not sure they would have been able to thoughtfully reflect on their self-efficacy throughout the semester which would have likely contributed to more disagreements in the results or missing data.

With respect to their equivalence, the treatment and control groups were not comparable on the basis of their Total SAT scores. The control group's scores were statistically significantly higher, even though their Math Section SAT scores were comparable to those of the treatment

group. These conclusions were based on the results of an Independent Samples Test (T-test). If the current study had a larger group of participants, then the Total SAT scores could be included as a control variable. Also, the treatment group had noticeably higher participation from students who were male, White, and/or from rural high schools than those who participated in the control group, while students who were female, Black, and/or from urban high schools had higher rates of participation in the control group. These differences in demographics were highlighted earlier in this chapter. However, it is important to reiterate that even though quantitative analysis did not find any statistical differences between the two groups, differences between them still exist and it is possible that these differences could have impacted the results. It is possible that larger numbers of participants could have helped to alleviate any differences between the two groups.

Another important limitation was instructor bias, which became increasingly evident in data analysis. Playing the role of instructor and researcher can be difficult to manage, and so, measures were put in place during data collection to mitigate this influence. Undergraduate student data collectors were used to collect all data. In addition, all data, with exception of the audio recorded interviews, were placed into sealed folders and not viewed until final semester grades were assigned. Once I, the instructor and researcher, could see the data and began data analysis, it was difficult to separate the two roles I held in relation to the participants. This bias was especially apparent while establishing inter-rater reliability. I, as the instructor, personally knew the students whereas my colleague did not. For example, I knew that Participant 2 appeared to have a big ego while in my classroom. My colleague did not know this about this participant when coding, and so the statement "I felt like it was too easy at a certain point. But then I got challenged" (Journal 2, p. 5) was one we discussed in detail and agreed on the Variable code. Another example is when I read through their Journals, I could hear their voices and see

172

their personalities emerge as I read their written words. I also knew, from interactions with students in the classroom, which students frequently used sarcasm, humor, or who liked to complain for sport more than others. My colleague might take a statement as true at face value, when I knew that the student had a character that was supportive of being joking or sarcastic. A second instance of when issues with bias arose were, again during data analysis, while I was working on the comparisons of data sources. It was difficult to separate the knowledge I have of each student and only consider the data. This is why I analyzed the assessments of agreement a total of four times to hopefully ensure the most accurate results. During these rounds, minor changes were made at times and noted. It is, however, important to acknowledge that my bias as an instructor is still present even though precautions were used.

In addition to bias affecting analysis, it is also important to indicate bias in teaching. Prior to the beginning of data collection, I made a commitment to teach the course as it was taught in Fall 2018. Overall, I feel that I upheld this commitment as much to my ability as possible. I adhered to the schedule that I laid out for students at the beginning of the semester. I also felt that my interactions with students were representative of how I interacted with students in the past, even though these are things I can never ensure because of lingering bias. This semester, as in the past, I always engaged students in conversations that were not necessarily focused on the course content to help foster the classroom community. My presentation of the course was also as consistent as possible with semesters in the past. There were, however, some events that were C4 Scholar Program related and the resolution of these events caused some change in what would have been a normal class period. For example, the drama that transpired between some of the students was a distraction, even for those who were not involved. It was distracting enough to the point that one of the Wednesday instructor workshops was dedicated to a behavioral

workshop. In that session, someone from the Office of Student Conduct came in and led the session. The way in which the C4 Scholar Program handled events like these would have been how they were handled with past cohorts, there just was never an occurrence to this extreme before.

Also related to instructor bias is the influence that I, the instructor, could have on the current research. In order to mitigate the influence that my role as the instructor could have on the results of the study, I took many precautions. These included not knowing the identities of any participants during the semester, having student data collectors collect the three data sources from participants, not looking at any data until final grades were submitted, and making sure that students were aware of each of the precautions that were taken. These precautions were in place so that the collected data was as authentic as possible and not influenced from my presence. It is, however, important to acknowledge that my bias as an instructor is always present no matter how many precautions are exercised to lessen it.

A final component of instructor bias was the issue of time. As an instructor and researcher of the current study, instructor journaling was difficult, and this was acknowledged prior to beginning data collection in Chapter 3. During my experience, it was challenging to make notes about every student each class period, and entries about every student did not always happen every day or even every week. The issue of time played a role here because there would be instances when I would not write about a specific student for some time because there were other events or issues that came to the forefront of a class session. The challenge of instructor journaling is best defined by the circumstantial nature of potentially documented instances because they may occur while I am with another student, they may occur outside of the

classroom, and/or they may be difficult to see since a student's self-beliefs are not a physical characteristic.

Another important limitation to discuss concerns the small number of statements in some of the qualitative data. Because of these small numbers and the intervals of percentages I assigned to each level of self-efficacy for data analysis, the number of coded statements could swing the ratings for students. As indicated in Chapter 3, statements were identified and coded appropriately as positive, negative, or variable. During inter-rater reliability, when adjustments to my coding were made, the addition or deletion of one code could change the rating assigned to a Journal, for example, from Moderately Positive to Very Positive, or for any other ratings that border a threshold as outlined in the coding schemes. However, after completion of establishing inter-rater reliability and ensuring that all qualitative data was coded with the agreed upon perspective that my colleague and I determined, I strongly feel that the ratings were representative of the participants in the study.

I would like to highlight a specific example to further illustrate this limitation with reference to a Journal for Participant 14. My colleague and I each assigned two positive statements to this student, but I coded four negative statements and he coded three negative statements. I rated this Journal as Somewhat Negative and he rated this Journal as Neutral. We were only off by one rating, but this example illustrates how the addition of one coded statement could affect the rating of a data piece.

APPENDICES

APPENDIX A: C4 Scholar Program

Description of the Evolution of the C4 Scholar Program and Development of Tools

C4 is a program for at-risk students who qualify for developmental placement in reading, English, and mathematics. If students choose to participate in this program, they save time by being placed into courses that are one level above their placement (based on ACT scores and high school GPA), they save money from not having to take non-credit bearing courses, and they are given extra supports that they normally would not receive in order to be successful. These supports include instructor-led workshops, a learning community of peers, and instructors who communicate across disciplines frequently throughout the year in order to holistically support student progress. A focus of our program is retention, which we hope to achieve through instilling accountability and professionalism in our students. After the pilot year (2016-2017), we incorporated the Oz Principle which focuses on these two ideas through above-the-line and below-the-line behaviors (Williams et al., 2020). The past two years (2017-2018 and 2018-2019) has been dedicated to incorporating these ideas. We hope to expand to another cohort of students and additional instructors in the next year or two.

A common thread that links the three courses are Enhanced Learning Points (or ELPs). These points are part of students' grades and are designed to encourage above the line behaviors and habits of successful students. Students can receive points for going to the tutoring center, coming to office hours, attending optional study sessions or workshops, revising their work, volunteering in class, etc. If a student did not engage in any of these activities, the best grade they could receive in any course (including mathematics) is a B. In addition to ELPs, I try to connect students' future careers and everyday lives to the course. In the past, this has been attempted by assigning a research paper for which students wrote about how mathematics was used in their chosen field, mathematics related to budgeting (car loans, mortgages, monthly budgets, banking), and occasional reflections that tie their behaviors in class to consequences of that same behavior in a professional setting (i.e. how would an employer react to…?). Some of these assignments have been successful in helping students see the connections between school and their future career or everyday lives, while others have not. For example, the research paper was not successful with the pilot group of students but would have perhaps been successful with a later cohort.

In addition, a foundation must be built in order to help students make steps toward more frequent above-the-line behaviors. In order to build this foundation, I offer many opportunities for success and gaining trust from the students throughout the semester. Traditionally, these students have not had the best experiences with mathematics, which I have learned from beginning of the year questionnaires and conversations with them. I begin the semester by helping students learn to become more comfortable with doing mathematics. I gain their trust by providing activities that are accessible, not rigid, and by encouraging collaboration with their peers. I work to lessen the fears and negative feelings students have learned to associate with mathematics, especially the dreaded story problems.

- We begin the first unit with the basics including order of operations, fractions, ratios/proportions, and percentages. Students are given the opportunity to engage with difficult, multi-step problems, connect to real world applications, and build their own problems that connect to their own interests. This unit was inspired by fears and gaps in skills that students in this program typically have.

- The second unit moves into linear equations, linear systems, exponents, and exponential equations. Students engage with budgeting for a vacation in which the number of days is unknown, but the variable costs are known. Students also engage with building/designing an amusement park and solving money related problems with banking, mortgages, and car loans. Ideas for these mini projects were inspired by resources made public and found online. This unit was inspired by my experiences with the curriculum of another course at FSU and the connections students could make to their everyday lives.
- The third unit expands into logarithmic equations, factoring, and quadratics where a variety of application-based problems are explored. This unit was inspired by my experiences with the curriculum of another course at FSU and engaging students with difficult problems, including story problems.
- Each unit provides a week to explore each new idea and ends with a week to connect ideas, review, and take an exam. Each unit is approximately 5 weeks in length.

The tools that I use take a variety of forms. I engage students with minimal lecture each class period and they often engage in conversations with their classmates about an activity that is provided. This activity is often given as a worksheet or prompts via PowerPoint that leads them through different ideas in order to improve their skills associated with the current unit. Sometimes, activities are lengthier explorations of material and can be followed by short group presentations. I like to occasionally provide students with opportunities to create their own problems and trade with another student to solve. Students are usually led through this process: identifying a related interest and variables, developing the situation and problem, writing a realistic problem to solve, and providing a solution.

The past three academic years with the C4 Scholar Program has been used to develop tools that I (and a co-collaborator at times) found to be effective based on student engagement (collaboration with each other, student attitudes, students staying on task) and student performance. The first year (2016-2017) involved a lot of trial and error. My co-instructor and I decided to explore the route of filling in the gaps that many students at this level may have. We began the year with activities about number sense – what numbers mean, different types of numbers, positive/negative numbers, operations, etc. We also gave a diagnostic exam, which was the previous year's final exam for one of my sections of the same course. This helped us to identify the directions to go for the rest of the semester. We followed this with work with fractions, ratios and percentages, and rational expressions/equations. The following spring semester focused on loans and money, logarithms and exponentials, and a short unit on basic statistics. Upon conclusion of this academic year, the instructors of C4 found and agreed that our population of students had deficits in the areas of accountability and professionalism. The English instructor of our group found research on the Oz Principle. We each read about it over the summer and felt it would be a great fit for our students.

The following academic year (2017-2018), we worked to incorporate accountability and professionalism into our courses. The math course during this second year followed a similar path, but the number sense unit was dropped and a focus on problem solving took over. Students were given a variety of application-based problems that focused on fractions, ratios/proportions, percentages, and quadratics. During the spring semester, I was on maternity leave for the first half. Upon returning, students were engaged with mastering basic skills via an online program Khan Academy outside of class to address any gaps in their content knowledge. They attended to budget-based mathematics (loans/interest, mortgages, car payments, etc.) and linear equations

during class time.

The last academic year (2018-2019), we decided to save students time and money, and merged the two semester algebra sequence down to one semester. I no longer taught the course with my co-instructor since he decided to take over the data aspect of the project. The current design of the course, with influences from materials used in the previous two years, materials I have pulled together from past courses, and online resources, is divided into three units and was previously described in this appendix. The decision to use specific materials was made with the following in mind: a focus on the material's ability to help students break down concepts they may have once had or still have trouble with, accessibility for a variety of skill levels, and the depth of exploration that encouraged students to collaborate with each other instead of a repetitive "plug and chug" mathematics environment.

I now discuss why I think these tools have a positive impact. As stated previously, many of these students have not had the best experiences with mathematics. These negative experiences have stemmed from a lack of success, a poor relationship with an instructor, negative feelings associated with the subject, among others. I try to implement interventions, activities, and build relationships with them that provide success and good experiences with mathematics. I feel that this has made differences for these students because of their statements and actions. I frequently hear comments such as "Your class goes by too quickly" and "We love coming to math class." According to the project coordinator, the positivity associated with the mathematics course by the C4 students has been consistent, unlike the English and reading courses. I also feel that the engagement level of students is much higher than what we could traditionally expect. It can be difficult, especially of freshmen, to expect high levels of participation because of the new, college environment and perhaps a lack of confidence with mathematics. The three years I have had a cohort of C4 students has proven to be vastly different. These students feel a sense of community, they trust each other and their instructor, and they have experienced success to boost their self-efficacy and confidence that they associate with mathematics. This is why I think there have been positive experiences for students and success for C4 in the past, and why I would like to formally measure the changes in self-efficacy that are experienced by this group of students with the current study.

APPENDIX B: Consent Forms
Consent Form for Students Participating in the
Tracking and Understanding At-Risk College Students' Self-Efficacy in an Intervention
Mathematics Course study

You are invited to take part in a research study of self-efficacy while you are enrolled in your first mathematics course at this university. The purpose of the study is to learn more about college student self-efficacy toward mathematics. For this study, *self-efficacy toward mathematics* is defined as how confident you (as a student) feel about your abilities to be successful in math. Self-efficacy is important to study because it can affect many things including your performance (grades) and motivation. In addition, this study could help future students of the C4 Scholar Program by learning more about the math course that you are enrolled in.

Please read this form carefully and ask any questions you may have before agreeing to take part in the study.

What the study is about: The purpose of this study is to learn about your self-efficacy toward mathematics and how it may change in your first math course.

What we will ask you to do: If you agree to participate in this study, you will fill out three 15-minute surveys, complete five 20-minute journal entries, and participate in two 30-45 minute interviews, with the possibility of two short follow-up interviews if clarifications are deemed necessary. These items will ask you about your relationship and experience with mathematics and associated factors that may influence your self-efficacy toward mathematics. With your permission, we would also like to audio-record the interviews.

Further, your high school GPA and ACT scores may be used in the research for this project, this consent form requests your permission to possibly request this information. In addition to these records, we request your consent to use your quiz and exam scores from this course (MATH 115) in which you are currently enrolled. The use of these scores in the research study would serve the purpose of seeing if any correlation between your scores on major assignments and self-efficacy result. Only the researcher will be able to identify you with these records. Any information that is made public will not be connected to your identity.

Please initial here _____ if you consent to having your high school GPA, ACT scores, quiz scores, and exam scores requested and used for this research study.

In addition to the researcher, you will also interact with two student data collectors. These students are past C4 Scholar Program participants.

Why your participation is important: Your participation in this study is important because we want to know how to help students have a good relationship with mathematics and understand how we can, if possible, increase students' levels of self-efficacy toward the subject. This study may not only help you and your classmates, but also future students who enroll in the C4 Scholar Program at this university.

Risks: There is minimal risk in that you may find some of the questions about your experiences to be sensitive.

Benefits: Your participation will earn a total of $40 in gift cards redeemable at the on-campus bookstore. You will receive a $20 gift card half-way through the semester. You will receive another $20 gift card at the end of the semester. These two gift cards together comprise the $40 participation incentive.

There is also the benefit that the responses you provide could be empowering to your ideas about yourself doing mathematics. Student self-efficacy toward mathematics is a very interesting issue and we hope to learn more about students in order to help them succeed in the subject and for the success of C4 Scholar students in the math course.

Your participation will be confidential. The records of this study will be kept private. In any sort of report we make public we will not include any information that will make it possible to identify you. Research records will be kept in a locked file; only the researchers will have access to the records. If we audio-record the interview, we will destroy the recording after it has been transcribed, which will be soon after its taping.

Since your instructor is the researcher of this study, your participation will also be kept confidential from your instructor until final grades have been submitted on December 17, 2019. Only the student data collector will know your identity up to that point in time. Your responses on surveys and journals will be kept in sealed envelopes until this date. The only data that your instructor will have access to are your de-identified interview audio-files and/or interview transcripts. The sole purpose of having access to these items is to determine whether a follow-up interview should take place after interviews 1 and 2.

Taking part is voluntary: Taking part in this study is completely voluntary. You may skip any questions that you do not want to answer. If you decide not to take part or to skip some of the questions, it will not affect your current or future relationship with Michigan State University or Ferris State University. If you decide to take part, you are free to withdraw at any time.

If you have questions: The graduate student conducting this study is Lauren Cavner Williams and the MSU faculty members are Dr. Jack Smith and Dr. Monica Karunakaran. Please ask any questions you have now. If you have questions later, you may contact Lauren at cavnerla@msu.edu. If you have any questions or concerns regarding your rights as a subject in this study, you may contact MSU's Institutional Review Board (IRB) Chairperson at 517-355-2180 or access their website at http://hrpp.msu.edu.

You will be given a copy of this form to keep for your records.

Statement of Consent: I have read the above information and have received answers to any questions I asked. I give consent for myself to take part in the study.

Name of Student (printed) _____

Student Signature _____ Date _____

In addition to agreeing to participate, I also consent to having the interviews audio recorded.

Your Signature _____ Date _____

Signature of person obtaining consent _____ Date _____

Printed name of person obtaining consent _____ Date_____

This consent form will be kept by the researcher for at least three years beyond the end of the study.

Consent Form for Students Participating in the
Tracking and Understanding At-Risk College Students' Self-Efficacy in an Intervention Mathematics Course study

You are invited to take part in a research study of self-efficacy while you are enrolled in your first mathematics course at this university. The purpose of the study is to learn more about college student self-efficacy toward mathematics. For this study, *self-efficacy toward mathematics* is defined as how confident you (as a student) feel about your abilities to be successful in math. Self-efficacy is important to study because it can affect many things including your performance (grades) and motivation.

Please read this form carefully and ask any questions you may have before agreeing to take part in the study.

What the study is about: The purpose of this study is to learn about your self-efficacy toward mathematics and how it may change in your first math course.

What we will ask you to do: If you agree to participate in this study, you will fill out two 15-minute surveys. This survey will ask you about your confidence to do various math tasks and complete various college courses that require mathematics. In addition, your high school GPA and ACT scores may be used in the research for this project, this consent form requests your permission to possibly request this information. These records will not be identifiable to you.

Please initial here _____ if you consent to having your high school GPA and ACT scores requested and used for this research study.

Why your participation is important: Your participation in this study is important because we want to know how to help students have a good relationship with mathematics and understand how we can, if possible, increase students' levels of self-efficacy toward the subject. This study may not only help you, but also future students at this university.

Risks: There is minimal risk in that you may find some of the questions about your experiences to be sensitive.

Benefits: Your participation will earn a $25 gift card redeemable at the on-campus bookstore. You will receive this gift card after you complete both surveys.

There is also the benefit that the responses you provide could be empowering to your ideas about yourself doing mathematics. Student self-efficacy toward mathematics is a very interesting issue and we hope to learn more about students in order to help them succeed in the subject.

Your participation will be confidential. The records of this study will be kept private. In any sort of report we make public we will not include any information that will make it possible to identify you. Research records will be kept in a locked file; only the researchers will have access to the records.

Taking part is voluntary: Taking part in this study is completely voluntary. You may skip any questions that you do not want to answer. If you decide not to take part or to skip some of the questions, it will not affect your current or future relationship with Michigan State University or Ferris State University. If you decide to take part, you are free to withdraw at any time.

If you have questions: The graduate student conducting this study is Lauren Cavner Williams and the MSU faculty members are Dr. Jack Smith and Dr. Monica Karunakaran. Please ask any questions you have now. If you have questions later, you may contact Lauren at cavnerla@msu.edu. If you have any questions or concerns regarding your rights as a subject in this study, you may contact MSU's Institutional Review Board (IRB) Chairperson at 517-355-2180 or access their website at http://hrpp.msu.edu.

You will be given a copy of this form to keep for your records.

Statement of Consent: I have read the above information and have received answers to any questions I asked. I give consent for myself to take part in the study.

Name of Student (printed) _____

Student Signature _____ Date _____

Signature of person obtaining consent _____ Date _____

Printed name of person obtaining consent _____ Date_____

This consent form will be kept by the researcher for at least three years beyond the end of the study.

APPENDIX C: Blank Data Tracker Sheet for Student Data Collectors

Table 30: C4 Student Data Tracker

Participant	Consent Complete	Survey 1 Complete	Journal 1 Complete	Interview 1 Complete	Journal 2 Complete	Survey 2 Complete	Journal 3 Complete	Journal 4 Complete	Interview 2 Complete	Journal 5 Complete	Survey 3 Complete

APPENDIX D: Description and Timeline of Classroom Drama
Taken from the Daily Instructor Journals

Week 7 – Tuesday
In class, from my observations, Participant 8 had been doing fine until yesterday. The class was particularly difficult to keep focused yesterday. Participant 8 (along with a few of his classmates) were quite distracted at the beginning of class. They eventually quieted down while I began the lesson, but when I provided group collaboration time on some practice problems, the activity picked up again. But, this time it seemed to radiate throughout the room. Many students were still focused, but others kept looking back at Participant 8 (and company) after looking at their phones. Towards the end of class (8 minutes to go), Participant 8 asked me if he could leave. I said no and asked why. He said that he was very angry about certain people in the room acting like high schoolers. He literally looked like he was about to blow up. I told him he could step outside for a minute. He stepped outside - I then heard a loud noise (I'm not sure if this was from him, but it sounded like he hit something), then he came back inside - seemingly "refreshed". I told him we could talk after class if he'd like, he agreed to. At the end of class, I said to everyone "I can tell that something is going on today. If anyone needs to talk with me about an issue, please feel free to stay behind and we can get it figured out." With this statement, Participants 3 and 9 stayed behind. Participant 8 had also come up, I told him that we can either wait for everyone to clear out of the room or we could go up to my office. He said okay, but then left. After everyone else left, Participants 3 and 9 started to speak with me about what was going on. They began by showing me a snapshot of a post that Participant 8 had made that was public. It was as the C4 coordinator described - a photo of a couple students in the classroom (during yesterday's 3-3:50pm class time) but from behind sitting at the tables, along with foul language in the posting - I do not remember the exact wording. Participant 9 also showed me messages between herself and Participant 8 where there was clearly a disagreement going on, again with foul language. Participants 3 and 9 said that this started to escalate on Monday (and possibly during homecoming over the weekend) when they decided they didn't want to hang out with Participant 8 (and others) as much anymore. This, they believe, is the source of the problem. From the messages I read off Participant 9's phone, this does appear to be consistent. Participant 3 was pretty upset, she teared up a couple times. I brought them up to the C4 coordinator's office yesterday after class to see what we should do next. When I brought them up, I looked for Participant 8 - to see if he was waiting for me - and he wasn't there. We (myself, Participant 3, and Participant 9), then, spoke with the C4 coordinator for about 15-20 minutes about what was going on.

Week 7 – Wednesday
It became apparent that Participant 2 was also involved. He emailed me before class and said he wasn't going to come today because of the drama. In lieu of yesterday's events, I decided to change groups up early to separate those involved in the drama from each other. Participant 8 was removed from the class for the rest of the week, the Office of Student Conduct made this decision. Class proceeded without interruption and is described in detail in the daily instructor journals.

Week 8 – Monday
Participant 8 transitioned back into the classroom today and he did great. I thought he was very respectful and focused throughout the lesson. I even got him on a plan to get caught up on what he missed from last week. He will take his missed quiz on Wednesday in the study session, then I will pass back everyone's quizzes when he is done. I spoke with Participants 3 and 9 before class today and they were fine moving forward – they had already seen him in their first class today. Participant 9 got an apology over the phone over the weekend from him, but Participant 3 had not. I will just make sure everyone is separated until the end of the semester. Overall, today went very well and we covered more than I hoped!

Week 10 – Wednesday
I noticed that Participant 2, Participant 8, and two non-participants were very distracted in the back of the class today, especially with a non-participant's phone. They were laughing. I asked them if they were alright, they said yeah. I asked them if they wanted to share what was so funny and they said no. They quieted down more for the rest of the class time. After class, another non-participant and Participant 5 stayed behind and they told me that the group was very tense in an earlier class today. They said that the drama has not quieted down and that people are still posting things about each other on social media. They said their instructor from the earlier class didn't do anything to stop or mediate it. So, I called the C4 coordinator right after class. I told him what was going on, who was involved, and he was not thrilled with the news. He said he would come down before the 4:30 class and talk to the four students I mentioned…. The C4 coordinator stayed for a few minutes and told me that there are other things going on beyond what we know. He said we should probably have someone from the Office of Student Conduct come in and lead a session.

Week 11 – Monday
Before I got to class, I received an email from the C4 Coordinator that said something happened between Participants 2 and 8, and involved campus police, over the weekend. After speaking with the C4 coordinator before class, I learned that Participant 2's girlfriend broke up with him and Participant 8 tried to move in on her. Participant 2 got mad, and left a threatening note under Participant 8's door. Participant 8 turned the note in to his RA, who got campus police involved.

When I got to the classroom, Participant 8 was one of the first people there. He went to sit in the back of the room. More people filed in over the next ten minutes. When Participant 2 and two other non-participants arrived, they were all quite stoic and went to the opposite side of the room. This is a drastic difference from before, as they all were friends. The C4 coordinator also came in just after everyone arrived. In my journal, I go on to explain how I tried to make the class period go as normally and smoothly as possible. I learned after class that Participant 2 did apologize to Participant 8 for sending him the note.

Week 11 – Tuesday
When I first arrived to class, most everyone was out in the hallway. Participant 7 was crying and being consoled by a friend. Participant 12 was looking at her phone pretty concerned, surrounded by Participants 6 and 14. I asked Participant 6 what was going on and he said a lot of stuff went down last night, he pointed toward Participant 8 (and mentioned Participant 2), and that stuff had been going on this afternoon. He said we can talk to Participant 8 about it because he is pretty

open. We all went into the classroom, and I got things organized and pulled up for the day's activities. When the C4 coordinator arrived, I took him out into the hallway and told him what I just observed. When we came back in, I noticed he spoke privately with Participant 7 for a second (he told me later that she had emailed him pretty upset yesterday and missed the study session because of it – he encouraged her to keep attending classes).

Week 11 – Wednesday Workshop (Behavioral Workshop)
During the workshop, we had a guest from the Office of Student Conduct come and do a workshop on professional relationships with colleagues. Students went along with the workshop for the most part. I do think it was helpful. I think it really got students thinking, especially the idea about empathy and everyone (mostly) agreeing to move forward for the rest of the semester. However, Participant 2 got up and left right away – I'm still not sure what that was about. But, we did not do anything related to math in this session. I just let our guest have the floor. Participants 15 and 19 didn't seem to take the workshop too seriously. I thought that Participants 6, 7, and 16 did. Participant 8 was a willing participant. Participant 10 didn't seem overly thrilled that this was happening in place of his math class. This speaks about his self-efficacy because he was disappointed that we weren't doing math! The C4 coordinator introduced our guest at the start and then I sat in the back and observed while it took place. Before the workshop, Participant 7 talked to the C4 coordinator outside – apparently there was another social media post by Participant 2 – so that could be why he got up and left because I believe that the C4 coordinator may have passed by him in the hallway. The timing would have been about right for that to happen.

Week 12 – Monday
Participant 2 and Participant 8 are sitting next to each other in class. It appears they are getting along again.

This seems to conclude the drama between students, that I could observe in class.

APPENDIX E: Sample Activities and Assignments

Sample Activities and Assignments in the Treatment Course

Activity 1: Given as a handout on week 1, day 1 of class.

There's More Than One Way
MATH 115 – C4 Project

Without using a calculator, find as many ways (minimum 3 ways) as you can to solve 18 x 5. Try using diagrams, pictures, unpack the numbers, etc. Get creative! We will share these later.

Activity 2: Given as a handout on week 4, day 2 of class.

MATH 115
Creating Percentage Problems

1. What is your favorite store? _____

2. What types of products do you like to browse or buy? _____

3. Based on your answers from #1 and #2, write a percentage mark-down problem.

4. Solve your problem from #3 here. Show your steps and circle your answer.

5. Based on your answers from #1 and #2, write a percentage mark-up problem.

6. Solve your problem from #5 here. Show your steps and circle your answer.

Students are then asked to rewrite their problems on the back of their paper and switch with a neighbor.

190

Activity 3: Given on PowerPoint slides for week 11, day 1 of class.

Slide 1:
You work for a relatively new, small firm that sells pet food. Your boss wants to put you in charge of the company's financial future and is hoping to set some money aside in case of emergency. Your boss wants to invest $5,000 into an account but isn't sure which one to choose. Here are the options:

- One that has 4.3% interest compounded quarterly.
- One that has 3.5% interest compounded continuously.
- One that has 5% interest compounded semiannually.

To help you know which is best, your boss wants you to advise her to choose the one that will get to $50,000 the fastest.

Which account will you advise your boss to invest in?

Note: Students have all of their tools from our unit on exponents and budgeting/money problems. They are given 10-15 minutes to come to a consensus with their group. This problem marks our transition from exponents to logarithms. They will realize that they are unable to solve the problem, although some students try alternative guess and check methods and come very close to the answer. After all groups give their answer and we hold a class discussion, I follow with…

Slide 2:
So, we need to know some tricks. This is where logarithms come in handy (because they "undo" exponentials). Which rule looks like it could help us?

- $\log_b AB = \log_b A + \log_b B$
- $\log_b \frac{A}{B} = \log_b A - \log_b B$
- $\log_b A^c = c \log_b A$

Note: We, as a group, explore and discuss these rules. Then, students return to their groups and try to apply it to the problem that was given at the beginning of class.

APPENDIX F: Treatment Course vs Other Sections

Details of Differences Between Treatment Course and Other Sections of Similar Course

Table 31: Content Differences

Treatment Course	Other Courses (MATH 110/115)

1. Spend more time on individual topics than 110/115
   o More time is given to number sense topics
2. Instructor covers major skill and content topics for the sequence, but some topics were cut
   o I cut factoring even though we use it in the quadratics section, they have seen it before and are encouraged to ask questions if they need help
   o I cut rational expressions/equations because of the degree paths that the majority of students are pursuing
3. Real-world examples often begin a lesson
	1. Some topics are given less time and attention than others, especially number sense topics at the beginning of the semester.
	2. A university list of course objectives is given to instructors as a basis to design their course. The 110/115 sequence was used as the "bank" of topics for the intervention course.
	3. Definitions and equations often begin a lesson, while an application problem may conclude a lesson.

Table 32: Structure Differences

Treatment Course	Other Courses (MATH 110/115)
1. Students get credit for sequence in one semester, less credits to pay for (3 credits, coded as MATH 115) 2. Two workshops (for ELP) are offered. Monday session is for general work/study and Wednesday session is for extra math help. o Workshop is led by instructor.	1. Students get credit for sequence in two semesters (4 credits, 3 credits) 2. If the course section has a required workshop (not all do), it is led by an older student

Table 33: Instructional Differences

Treatment Course	Other Courses (MATH 110/115)
1. Instruction consists of minimal lecture but with a supportive introduction of topic, many opportunities for exploration/group work, instructor acts more as a facilitator rather than a lecturer in order to open discussion. 2. This opens space for me to move	1. Traditional instruction – consists of a lecture format, minimal exploration/group work, instructor acts as a lecturer. 2. Fewer instances of encouraged interaction between students and instructor or other students

Table 33 (cont'd)

around the room to check-in with groups and talk with students	3. Instructor can view every submitted problem or assignment if s/he goes into the online portal and clicks through each problem for each student.
3. Instructor sees every submitted assignment and will know where students struggle with material. Instructor encourages students to ask questions upon returned assignments and reminders of this are made regularly	

Table 34: Assessment Differences

Treatment Course	Other Courses (MATH 110/115)

1. Paper/pencil homework (counts toward student grade)
   o Prompt, handwritten feedback from instructor
   o All work, including a final solution is required from students. An indication of where an error was made is noted by the instructor. But correct answer is not provided so students can complete revisions.
2. In class quizzes and tests are given but are only weighted to be about 50% of the entire final grade
3. Revisions on all graded items are allowed except the final exam
   o When students revise their assignments, they receive an average grade between the original new scores. For example, if the student received a 60% and revised to an 80%, their final score would be 70%. If a student wants to revise an assignment, they must submit the original assignment with their revision within one week.
4. ELPs – created for C4 Scholar Program, encourages non-required student/instructor interactions and proactive behaviors
	1. Roughly 80% of all sections, per the departmental secretary, use online homework (counts toward student grade)
	   o Immediate feedback, not from instructor
	   o Only answers are submitted and whole problems are marked wrong, students do not know where they made an error and are not "required" to write work out. Correct answer may be provided, and students are given the option to view an example or try a new problem.
	2. In class quizzes and tests are given and are usually weighted as a significant majority of the grade (i.e. other sections the current instructor taught were worth 75% of the grade)
	3. No revisions are allowed
	4. No ELPs – minimal encouragement of student/instructor interaction and proactive behaviors

APPENDIX G: Journal Prompts

Journal Prompts 1 – 5

Journal 1 – Math Autobiography
- This journal is meant to document your story as a learner of math, past and present. Write about your experiences with math. Think of this as you are telling your story.
- *The questions below can help you with this free-write. You may use all, some, or none of these in your response. You have approximately 20 minutes, please use all the allotted time. If you need more time, a convenient space will be sought and provided for you (if necessary). Your response has no weight on your grade and will not be read by the instructor until final grades have been submitted.*
    - What comes to mind when you think of math?
    - Do successes or struggles stick out in your mind? Any situations in particular?
    - What was math like for you in elementary school? Middle school? High school?
    - Do you feel that you have support from others in the subject? If so, who? Please explain whether you answered yes or no.
    - Do you find math useful? This may not be just a simple yes/no answer. Please explain and provide examples.
    - Any closing thoughts?

Journal 2 – Unit 1
- What are your thoughts on Unit 1?
- *The questions below can help you with this free-write. You may use all, some, or none of these in your response. You have approximately 20 minutes, please use all the allotted time. If you need more time, a convenient space will be sought and provided for you (if necessary). Your response has no weight on your grade and will not be read by the instructor until final grades have been submitted.*
    - What do you think about the assignments and activities in this unit?
        - Were they helpful and/or useful? How?
    - Any successes or struggles that you experienced?
    - Any part of the material you feel confident about?
    - Any part of the material you still have questions about?

Journal 3 – Unit 2
- How did you feel about your Exam 1 score? Was the result what you expected? How did it make you feel about yourself and how you think about math?
- What are your thoughts on Unit 2?
- *The questions below can help you with this free-write. You may use all, some, or none of these in your response. You have approximately 20 minutes, please use all the allotted time. If you need more time, a convenient space will be sought and provided for you (if necessary). Your response has no weight on your grade and will not be read by the instructor until final grades have been submitted.*
    - What do you think about the assignments and activities in this unit?
        - Were they helpful and/or useful? How?
    - Any successes or struggles that you experienced?

- o Any part of the material you feel confident about?
- o Any part of the material you still have questions about?

Journal 4 – Unit 3
- How did you feel about your Exam 2 score? Was the result what you expected? How did it make you feel about yourself and how you think about math?
- What are your thoughts on Unit 3?
- *The questions below can help you with this free-write. You may use all, some, or none of these in your response. You have approximately 20 minutes, please use all the allotted time. If you need more time, a convenient space will be sought and provided for you (if necessary). Your response has no weight on your grade and will not be read by the instructor until final grades have been submitted.*
  - o What do you think about the assignments and activities in this unit?
    - ▪ Were they helpful and/or useful? How?
  - o Any successes or struggles that you experienced?
  - o Any part of the material you feel confident about?
  - o Any part of the material you still have questions about?

Journal 5 – Reflection on the Semester
- How did you feel about your Exam 3 score? Was the result what you expected? How did it make you feel about yourself and how you think about math?
- Write about your experience this semester in the math course.
- *The questions below can help you with this free-write. You may use all, some, or none of these in your response. You have approximately 20 minutes, please use all the allotted time. If you need more time, a convenient space will be sought and provided for you (if necessary). Your response has no weight on your grade and will not be read by the instructor until final grades have been submitted.*
  - o What was your overall experience with the math course?
  - o Do you feel any differently about or towards mathematics than in the past?
  - o Are there any significant successes or struggles that stick out in your mind?

APPENDIX H: Interview Protocols

Interview 1a Protocol
Thinking about the past.

*Begin recording now. Set device to DO NOT DISTURB, or recording will stop.

Before starting: "*This interview will help us to understand how you think about yourself and math. The first questions are going to be about you, in general, and help us start the interview.*"

1. How would you describe (*probe these questions, ask why or can you explain further*)…
   a. yourself?
   b. yourself as a student? (study habits, academic history, goals)
   c. your relationships with friends and family?
   d. your life in general, in and out of school?

2. Where did you attend high school (school name and city)?
   a. What was the last math class you completed at your school?
   b. When did you take that class?
   c. How did you feel about this class?

3. Describe how you currently feel about math.

Say: "*The rest of the questions will be about your experiences before college.*"
4. (*Mastery Experiences*) Before you came to college…
   a. Do you feel you have been successful with math?
      i. How do you define "being successful"?
   b. What struggles, if any, have you experienced with math?
   c. Are there specific situations that you think of where you struggled or felt successful?

5. (*Verbal Persuasion*) Did you talk with teachers, friends, parents, or other family about these successes or struggles?
   a. What were some of these discussions like?
   b. How have these conversations influenced you?

6. (*Vicarious Experiences*) Have there been times before you came to college that you observed…
   a. A classmate having success with math.
   b. A classmate struggling with math.
   c. Do you think these observations influence how you feel about math?
      i. If needed, help them refer back to specific examples from parts a, b

7. (*Emotional Arousal*) Before you came to college, how did you feel when you walked into math class if it was a typical class day?

8. (_Emotional Arousal_) Before you came to college, how did feel when you walked into math class if it was a test or quiz day? Explain.

*Skip #9 if running low on time

9. Do you feel that your feelings about math been influenced by
   a. who you are as a person?
   b. who you are as a student?
   c. your relationships with others?
   d. your life in general?

10. Have your feelings about math changed over time?
    a. If yes, please explain.
       i. What do you believe has contributed to these changes?

*for this question, look for specific mention of time or grade levels in school

Say: "_Is there anything else that you wanted to talk about or anything that came to mind?_"

Interview 2a Protocol
Thinking about this semester.

*Begin recording now. Set device to DO NOT DISTURB, or recording will stop.

Before starting: "*This interview will help us to understand how you think about yourself and math after taking your current college math course. The first questions are going to be about you, in general, and help us start the interview.*"

1. How would you describe…
    a. yourself?
    b. yourself as a student? (study habits, academic history, goals)
    c. your relationships with friends and family?
    d. your life in general, in and out of school?

2. Describe how you currently feel about math.

3. (Mastery Experiences) Thinking about this semester…
    a. Do you feel you have been successful with math?
        i. How do you define "being successful"?
    b. What struggles, if any, have you experienced with math?
    c. Are there specific situations that you think of where you struggled or felt successful?

4. (Verbal Persuasion) Did you ever talk with teachers, friends, parents, or other family about these successes or struggles?
    a. What were some of these discussions like?
    b. How have these conversations influenced you?

5. (Vicarious Experiences) Have there been times this semester that you observed…
    a. A classmate having success with math.
    b. A classmate struggling with math.
    c. Do you think these observations influence how you feel about math?
        i. If needed, help them refer back to specific examples from parts a, b

6. (Emotional Arousal) How did you feel when you walked into math class this semester if it was a typical class day? Explain.
    a. Did this change at all throughout the semester?

7. (Emotional Arousal) How did you feel when you walked into math class this semester if it was a test or quiz day? Explain.
    a. Did this change at all throughout the semester?

8. (Emotional Arousal) How did you feel when you got a math assignment this semester? Explain?
    a. Did this change at all throughout the semester?

b. How would you compare your feelings to how you felt if you got a math assignment in high school?

*Skip #9 if running low on time
11. Do you feel that your feelings about math been influenced by
   a. who you are as a person?
   b. who you are as a student?
   c. your relationships with others?
   d. your life in general?

10. Have your feelings about math changed over time this semester?
   a. If yes, please explain.
      i. What do you believe has contributed to these changes?

11. Did any non-academic events or distractions affect your college work this semester?
   a. If yes, Please explain.
   Note: If anyone asks, "What do you mean?", you can refer to "social issues between the C4 classmates".
   b. Did any such events affect your work in the math class?
   c. Did anything else affect your experience besides these events?

Say: "*Is there anything else that you wanted to talk about or anything that came to mind?*"

APPENDIX I: Student Data Collector Preparation

Preparing the Student Data Collector to Interview Participants

The student data collector joined the researcher during the pilot interviews prior to the beginning of actual data collection. This occurred in August 2019. The purpose of this was to allow the student to observe the structure of the interviews and how questions are asked, learn when to probe participants for more information, ask questions about the process, and practice asking participants some of the interview questions.

Three pilot interviews were conducted. The student data collector observed the entire first pilot interview as I asked the questions to the volunteer. The student data collector observed the first half of the second pilot interview and asked questions during the second half. The student data collector led the entire third pilot interview.

After each pilot interview, the student data collector and I discussed the flow of the interview with the volunteer. We asked them what their experience was with the interview and their feedback. We also talked about the interview after the volunteer left the conversation. All of these conversations helped to shape the interview protocol used for the study.

The student data collector met with me before, during, and after the 1a interviews. He also met with me before, during, and after the 2a interviews. This allowed the student to ask any further questions about data collection and allowed me to discuss anything I was noticing with the interview technique used in the interviews that had been completed. In particular, after we met during the 1a interviews, I noticed there to be a significant improvement in the student data collector's interviewing skills. Before he and I met, he had interviewed 7 participants. Of those 7, I asked that he follow-up with 4 of them. Of the remaining interviews after our meeting, I only asked that he follow-up with 1. Because my student data collector's interviewing skills were improving, I wanted him to follow up with more earlier interviews so that they are of the same quality as those that followed our meeting mid-way through interview 1a. If I didn't feel that I got a good "picture" of each student following their first interview (in relation to their self-efficacy), or if they didn't explain an answer or if they answered questions incorrectly (i.e. about C4 and not prior to college), then I recommended that we follow-up with interview 1b. For each of the 5 that I chose, I outlined personalized questions based on what was said in interview 1a – I even quoted questions and answers to help jog their memory.

APPENDIX J: Inter-Rater Reliability

Details from Inter-rater Reliability

The first is the participants' use of the word *help* in their journals. I coded these as positive statements if the participant explicitly gave details and connected it to him/herself. For example, "the assignments and activities she made us do is very helpful. They helped me learn the formulas and make them easier to understand the problems" (Journal 2, p. 15). Statements like "the assignments were helpful" with no reasoning or connection to him/herself were not coded. After consulting with my advisors, it is important to note that the initial feelings of the word "help" are generally positive from a reader's point of view, especially if the participant explained how/why something/someone was helpful. To determine the number of statements in which this issue arose, I reviewed each journal and interview for all participants. The number of instances detailed below are for the number of times that a student indicated that something helped AND provided detail about how it helped him/herself.

- Journal 1: 4 out of 118 coded statements = 3%
- Journal 2: 14 out of 114 coded statements = 12%
- Journal 3: 6 out of 144 coded statements = 4%
- Journal 4: 7 out of 122 coded statements = 6%
- Journal 5: 3 out of 135 coded statements = 2%
- Interview 1: 37 out of 618 coded statements = 6%
- Interview 2: 39 out of 761 coded statements = 5%

As a result, an average of 5.4% of coded statements would be affected if these statements were not coded positive. Since the addition or deletion of one statement could affect the ratings of a data piece, especially for journals, I felt that not including these statements could have a negative effect on the validity of the final overall codes. Therefore, the decision to code these statements as positive was confirmed.

The second is when participants used the phrase "I have questions about…" a mathematical topic. Sometimes, one of us would code the statement as negative, while the other would not code it at all because just having questions is not a negative thing and questions were encouraged in the course. After speaking with my advisors and colleague, it was determined that these statements are open for interpretation and we were still making these interpretations as we coded each journal. There were, however, a couple of clear notes to make about this issue. If a participant simply stated s/he had questions, then this was not coded. But, if the participant said they were really struggling and had questions, then that was coded as negative. Further, if the participant noted something positive and negative, then it was coded as variable. Like the issue that was previously discussed about the word *help*, I decided to review each journal and interview for all participants to see how many instances this arose in the data. It is important to note that the participants had many instances in the interviews about not being afraid to ask questions. These were not included in the count below because the student did not say "I still have questions about fractions" or something similar. Those were different kinds of statements and coded accordingly. Here are the number of times that statements about having specific questions about the material arose.

- Journal 1: 0 out of 118 coded statements = 0%
- Journal 2: 4 out of 114 coded statements = 3.5%
- Journal 3: 4 out of 144 coded statements = 3%

- Journal 4: 2 out of 122 coded statements = 2%
- Journal 5: 0 out of 135 coded statements = 0%
- Interview 1: 0 out of 618 coded statements = 0%
- Interview 2: 0 out of 761 coded statements = 0%

As a result, an average of 1.2% of coded statements would be affected and so the decision to allow this issue to be open to interpretation depending on how the student presented his/her questions would not affect much, if any, of the data analysis.

Summary. In addition to the codebook in Table 3, here are the issues that arose in our discussions, along with decisions on what code to apply.

Round 1 Changes/Notes:
1. Do not code statements that simply say "I have questions…" A statement like this does not indicate whether something is good or bad or how they feel. If there is more to the statement, then that can be assessed appropriately.
2. Statements about the instructor or class can be coded if there is an emphasis on how the individual participant is being affected positively or negatively.
   - The focus must be on <u>that student</u>, though.
3. Variable statements are single statements that show some time has passed and up/down patterns are present.
   - i.e. "I struggled at first, but then…good things happened" (example).

Round 2 Notes – No Changes Necessary:
- If participants mention that something or someone was "helpful", that invokes a positive feeling. But, for us to code the statement as positive, the participant must do one of the following (a) specifically connect a reason for how it was helpful to themselves in math or (b) make a connection to him/herself. If they do not do this, then we do not code it.
   o For example, "the assignments and activities she made us do is very helpful. They helped me learn the formulas and make them easier to understand the problems" will receive a positive code (Journal 2, p. 15).
   o In comparison, statements like "the assignments were helpful" with no reasoning or connection to him/herself are not coded.
- Statements when participants say something is "okay" are coded as variable. For example, "Unit 3 is going okay".
- Statements when participants say "I still have questions about…" can be open for interpretation when coding. If they include something that indicates what they have questions about is causing them to struggle, then we code it as negative. If it is a simple statement that s/he has questions about a topic with no other information, it should not be coded.
- If a participant indicates change across some time (i.e. improvement or worsening) about a particular idea in a statement, then that is coded as variable.
- Positive/Negative statements and reasons are included in the calculation to determine an rating.
- One of us splitting statements and the other did not – we still agreed on the rating of that statement. There was no question there. Only a question on the division. Instances where this occurred…

  o One of us coded 1 positive statement and the other split into 2 positive statements.
  o One of us coded 1 negative statement and the other split into 2 negative statements.
  o One of us coded 1 positive and 1 negative statement and the other combined into 1 variable statement.
    ▪ A specific example of this is when a participant would write about how their exam score came out better (or worse) than they thought it would. We decided this should be coded as variable.

Round 3 Changes
- The only new issue that arose, which was one in an extremely specific context, was in relation to participants' expectations versus reality of an exam score that they received back.
  - For example, participants would express self-doubt followed by excitement because they did not think they were going to perform as well as they did on an exam (and vice versa). My colleague would code this as Variable, and I would, code the self-doubt portion as negative and the excitement portion as positive.
  - Ultimately, we decided to code these as Variable because it best captures the change a participant undergoes.

Table 35: Round 2 for Second Half of Journal 2 – Final Codes

Participant	My Rating	Colleague's Rating	Agreed Upon Final Count and Rating
Participant 6	Moderately positive (1-5-0)	Moderately positive (1-3-0)	**Moderately positive (1-5-0)**
Participant 14	Neutral (4-4-0)	Neutral (5-7-0)	**Neutral *3 instances of splitting**
Participant 11	Very positive (0-4-0)	Very positive (0-7-0)	**Very Positive (0-6-0) *1 instance of splitting**
Participant 12	Moderately positive (2-6-1)	Somewhat positive (2-5-1)	**Somewhat Positive (2-5-2)**
Participant 8	Somewhat negative (5-3-0)	Somewhat negative (5-3-0)	**Somewhat negative (5-3-0)**
Participant 7	Moderately positive (1-3-0)	Moderately positive (1-5-0)	**Moderately positive (1-5-0)**
Participant 5	Moderately positive (1-4-1)	Very positive (1-9-1)	**Moderately positive (1-5-1) *2 instances of splitting**
Participant 13	Very positive (1-6-0)	Moderately positive (1-4-0)	**Very positive (1-7-0) *2 instances of splitting**
Participant 3	Neutral (3-3-1)	Neutral (2-3-0)	**Neutral (3-3-1) *2 instances of splitting**

Table 36: Round 2 for Second Half of Journal 2 – If Splitting Matters

Participant	Agreed Upon Final Rating if We Agreed to Split	Agreed Upon Final Rating if We Agreed to Combine	Is there an effect of choosing to split?
Participant 6	Moderately positive No instances of splitting		N/A
Participant 14	Neutral *3 instances of splitting	Neutral *3 instances of combining	No
Participant 11	Very Positive *1 instance of splitting	Very Positive *1 instance of combining	No
Participant 12	Somewhat Positive No instances of splitting		N/A
Participant 8	Somewhat negative No instances of splitting		N/A
Participant 7	Moderately positive No instances of splitting		N/A
Participant 5	Very Positive *2 instances of splitting	Moderately Positive *2 instances of combining	Yes, by 1 category
Participant 13	Very positive *2 instances of splitting	Moderately Positive *2 instances of combining	Yes, by 1 category
Participant 3	Neutral *2 instances of splitting	Neutral *2 instances of combining	No

Table 37: Round 3 for Journal 4 – Final Codes

Participant	My Rating	Colleague's Rating	Agreed Upon Final Count and Rating
Participant 4	Somewhat negative (5-3-0)	Neutral (3-4-1)	**Neutral** **(4-3-1)** ***1 instance of splitting**
Participant 6	Moderately positive (1-5-0)	Very Positive (0-4-2)	**Moderately Positive** **(1-5-1)** ***1 instance of splitting**
Participant 14	Somewhat negative (4-2-1)	Neutral (3-2-2)	**Moderately Negative** **(4-1-2)**
Participant 16	Neutral (3-3-0)	Moderately Positive (1-5-1)	**Neutral** **(4-4-0)**
Participant 2	Neutral (3-4-1)	Somewhat Negative (2-1-1)	**Neutral** **(3-4-1)** ***1 instance of splitting**
Participant 11	Very positive (0-5-0)	Very Positive (0-7-0)	**Very Positive** **(0-6-0)** ***1 instance of splitting**
Participant 18	Moderately positive (1-5-0)	Very Positive (0-3-2)	**Moderately Positive** **(1-4-1)** ***1 instance of splitting**
Participant 9	Neutral (4-4-1)	Neutral (3-4-2)	**Neutral** **(3-4-2)**
Participant 12	Neutral (3-4-0)	Moderately Positive (1-4-4)	**Moderately Positive** **(1-4-4)**

Table 37 (cont'd)

Participant 8	Very negative (4-0-1)	Neutral (3-2-1)	**Moderately Negative** ***we discussed this student in detail**
Participant 1	Very positive (0-5-0)	Very Positive (0-7-0)	**Very Positive** **(0-7-0)** ***1 instance of splitting**
Participant 17	Somewhat negative (2-1-0)	Very Negative (2-0-2)	**Very Negative** **(3-0-2)**
Participant 10	Very positive (0-6-0)	Very Positive (0-7-0)	**Very Positive** **(0-7-0)**
Participant 7	Somewhat negative (4-2-1)	Somewhat Negative (4-2-2)	**Somewhat Negative** **(4-2-2)** ***2 instances of splitting**
Participant 19	Moderately positive (2-6-0)	Somewhat Positive (2-5-0)	**Moderately Positive** **Agreed initially because my colleague thought it should be Moderately.** ***1 instance of splitting**
Participant 5	Very positive (0-7-0)	Very Positive (0-8-0)	**Very Positive** **(0-8-0)**
Participant 13	Very positive (1-6-1)	Very Positive (1-6-1)	**Very Positive** **(1-6-1)**
Participant 3	Somewhat negative (2-1-1)	Somewhat Positive (1-2-0)	**Somewhat Negative** **(2-1-1)** ***1 instance of splitting**
Participant 15	Somewhat Positive (2-4-0)	Somewhat Positive (1-2-3)	**Somewhat Positive** **(2-4-2)**

APPENDIX K: Comparisons Between the Data Sources

Table 38: MSES and Interviews

Participant	Interview 1	Interview 2	MSES 1	MSES 2	MSES 3	Level of Agreement	Reason
1	Neutral	Moderately Positive	4.15	4.65	7.59	Strongly Agree	Exact match
2	Very positive	Very Positive	7.82	8.15	8.36	Strongly Agree	Exact match
3	Somewhat Negative	Neutral	5.53	5.61	6.25	Somewhat Agree	Same trend
4	Neutral	Somewhat Positive	5.32	6.82	8.03	Strongly Agree	Same trend, close match
5	Neutral	Very Positive	4.97	5.53	5.65	Somewhat Disagree	Same start, opposite trend
6	Neutral	Somewhat Positive	4.15	5.50	6.03	Strongly Agree	Same trend, close match
7	Moderately Positive	Somewhat Positive	4.38	4.85	8.12	Somewhat Disagree	Opposite trend
8	Somewhat Negative	Neutral	2.79	5.12	5.35	Strongly Agree	Exact match
9	Somewhat Negative	Moderately Positive	4.18	5.44	6.12	Strongly Agree	Same trend, close match
10	Somewhat Positive	Very Positive	4.97	4.53	7.53	Strongly Agree	Same trend, close match
11	Moderately Positive	Very Positive	4.85	6.32	6.50	Strongly Agree	Same trend
12	Somewhat Negative	Somewhat Positive	3.64	5.18	6.26	Strongly Agree	Same trend, close match
13	Neutral	Somewhat Positive	2.29	4.26	6.91	Strongly Agree	Same trend, close match
14	Moderately Negative	Moderately Negative	3.71	3.82	4.30	Somewhat Agree	Same trend, close match
15	Neutral	Neutral	6.09	6.53	7.71	Somewhat Disagree	Opposite trend and one is much higher
16	Moderately Negative	Neutral	5.85	7.71	7.44	Somewhat Agree	Same trend
17	Moderately Positive	Neutral	6.50	6.82	7.88	Strongly Disagree	Opposite trend
18	Somewhat Negative	Somewhat Positive	4.26	6.79	6.94	Strongly Agree	Same trend, close match
19	Moderately Negative	Somewhat Positive	5.59	6.24	7.41	Strongly Agree	Same trend, close match

Table 39: MSES and Journals

Participant	Journal 1	Journal 2	Journal 3	Journal 4	Journal 5	MSES 1	MSES 2	MSES 3	Level of Agreement	Reason
1	Somewhat negative	Neutral	Very positive	Very positive	Moderately positive	4.15	4.65	7.59	Strongly Agree	Same trend
2	Very positive	Very positive	Moderately Positive	Neutral	Moderately negative	7.82	8.15	8.36	Strongly Disagree	Opposite trend
3	Neutral	Neutral	Neutral	Somewhat negative	Neutral	5.53	5.61	6.25	Somewhat Agree	Exact match until end
4	Somewhat positive	Moderately positive	Moderately positive	Neutral	Neutral	5.32	6.82	8.03	Strongly Disagree	Opposite trend
5	Neutral	Moderately positive	Very positive	Very positive	Very positive	4.97	5.53	5.65	Strongly Disagree	Opposite trend
6	Moderately positive	Moderately positive	Moderately positive	Moderately positive	Moderately positive	4.15	5.50	6.03	Somewhat Agree	Similar trend & end point
7	Neutral	Moderately positive	Neutral	Somewhat negative	Neutral	4.38	4.85	8.12	Strongly Disagree	Opposite trend
8	Moderately negative	Somewhat negative	Neutral	Moderately negative	Somewhat Positive	2.79	5.12	5.35	Strongly Agree	Same trend
9	Somewhat negative	Somewhat positive	Moderately Positive	Neutral	Somewhat Positive	4.18	5.44	6.12	Somewhat Agree	Similar trend
10	Neutral	Very positive	Very positive	Very positive	Very positive	4.97	4.53	7.53	Somewhat Agree	Similar trend
11	Neutral	Very positive	Somewhat positive	Very positive	Very positive	4.85	6.32	6.50	Strongly Agree	Same trend
12	Moderately negative	Somewhat positive	Somewhat positive	Moderately Positive	Moderately positive	3.64	5.18	6.26	Strongly Agree	Same trend
13	Moderately negative	Very positive	Somewhat positive	Very positive	Very positive	2.29	4.26	6.91	Somewhat Agree	Similar trend
14	Very negative	Neutral	Neutral	Moderately negative	Somewhat negative	3.71	3.82	4.30	Somewhat Agree	Similar trend
15	Neutral	Moderately negative	Moderately positive	Somewhat Positive	Somewhat negative	6.09	6.53	7.71	Strongly Disagree	Opposite trend
16	Moderately negative	Neutral	Somewhat Positive	Neutral	Neutral	5.85	7.71	7.44	Somewhat Agree	Similar trend
17	Neutral	Somewhat positive	Very positive	Very negative	Somewhat positive	6.50	6.82	7.88	Somewhat Disagree	Opposite trend
18	Somewhat negative	Somewhat positive	Somewhat positive	Moderately positive	Moderately positive	4.26	6.79	6.94	Somewhat Agree	Same trend
19	Neutral	Neutral	Somewhat positive	Moderately positive	Very positive	5.59	6.24	7.41	Strongly Agree	Same trend, close match

Table 40: MSES and Course Assessments

Participant	Exams	MSES 1	MSES 2	MSES 3	Level of Agreement	Reason
1	97, 99, 93, 95	4.15	4.65	7.59	Strongly Agree	Strong scores support increase
2	99, 68, 74, 58	7.82	8.15	8.36	Strongly Disagree	Declining scores do not support
3	78, 58, 91, 86	5.53	5.61	6.25	Strongly Agree	Same trend
4	96.5, 96, 70, 93	5.32	6.82	8.03	Somewhat Agree	Fairly consistent scores support increase
5	100, 96, 100, 96	4.97	5.53	5.65	Somewhat Disagree	Strong scores do not support consistency
6	96, 81, 84, 95	4.15	5.50	6.03	Strongly Agree	Fairly consistent scores support slight increase
7	87, 79, 85, 92	4.38	4.85	8.12	Somewhat Disagree	Consistent scores do not support sudden jump
8	91, 91, 82, 96	2.79	5.12	5.35	Somewhat Agree	Fairly consistent scores support increase
9	97, 79, 81, 91	4.18	5.44	6.12	Somewhat Agree	Fairly consistent scores support increase
10	94.5, 94.5, 95, 95	4.97	4.53	7.53	Strongly Agree	Strong scores support increase
11	96, 98, 99, 95	4.85	6.32	6.50	Strongly Agree	Strong scores support increase
12	97, 94, 95, 93	3.64	5.18	6.26	Strongly Agree	Strong scores support increase
13	100, 95, 87, 93	2.29	4.26	6.91	Strongly Agree	Strong scores support increase
14	85, 81, 55, 62	3.71	3.82	4.30	Somewhat Agree	Inconsistent and some weak scores support consistency
15	77, 73, 32, 66	6.09	6.53	7.71	Strongly Disagree	Declining scores do not support increase
16	95, 80, 55, 87	5.85	7.71	7.44	Somewhat Disagree	Declining scores do not support increase
17	96, 63, 66, 73	6.50	6.82	7.88	Strongly Disagree	Declining scores do not support increase
18	82, 69, 86, 89	4.26	6.79	6.94	Strongly Agree	Consistent scores support increase
19	93, 91, 89, 85	5.59	6.24	7.41	Strongly Agree	Strong scores support increase

Table 41: MSES and Instructor Perspective

Participant	Instructor Perspective (Week 5, Week 10, Week 16)			MSES 1	MSES 2	MSES 3	Level of Agreement	Reason (Week 5 is between MSES 1-2)
1	Mostly Positive	Mostly Positive	Very Positive	4.15	4.65	7.59	Somewhat Agree	Similar trend
2	Very Positive	Mixed – Positive	Mixed-Negative	7.82	8.15	8.36	Strongly Disagree	Opposite trend
3	Very Negative	Mostly Negative	Mostly Positive	5.53	5.61	6.25	Somewhat Agree	Similar trend
4	Mostly Positive	Mostly Positive	Mixed	5.32	6.82	8.03	Somewhat Agree	Same trend except end
5	Mixed-Positive	Mixed-Positive	Very Positive	4.97	5.53	5.65	Somewhat Disagree	Opposite trend
6	Mixed	Very Positive	Very Positive	4.15	5.50	6.03	Strongly Agree	Same trend
7	Mostly Positive	Mixed	Mixed-Negative	4.38	4.85	8.12	Somewhat Disagree	Very opposite end
8	Mixed	Mixed-Positive	Mixed-Negative	2.79	5.12	5.35	Strongly Agree	Same trend
9	Mixed	Very Positive	Mostly Positive	4.18	5.44	6.12	Somewhat Agree	Similar trend
10	Mostly Positive	Mostly Positive	Very Positive	4.97	4.53	7.53	Somewhat Agree	Similar trend
11	Mostly Positive	Mostly Positive	Mostly Positive	4.85	6.32	6.50	Strongly Agree	Same trend
12	Mixed	Mixed	Mixed-Positive	3.64	5.18	6.26	Somewhat Agree	Similar trend
13	Mixed	Mixed-Positive	Mixed	2.29	4.26	6.91	Somewhat Agree	Similar trend
14	Very Negative	Mixed-Positive	Very Negative	3.71	3.82	4.30	Somewhat Agree	Similar trend
15	Mostly Positive	Mixed-Negative	Mostly Negative	6.09	6.53	7.71	Strongly Disagree	Opposite trend
16	Mixed	Mixed-Negative	Mixed-Negative	5.85	7.71	7.44	Somewhat Disagree	Opposite trend
17	Mixed	Mostly Negative	Mostly Negative	6.50	6.82	7.88	Strongly Disagree	Opposite trend
18	Mostly Negative	Mixed	Mostly Positive	4.26	6.79	6.94	Somewhat Agree	Same trend
19	Mixed	Mixed-Positive	Mixed-Positive	5.59	6.24	7.41	Somewhat Agree	Similar trend

Table 42: Interviews and Journals

Participant	Journal 1	Journal 2	Journal 3	Journal 4	Journal 5	Interview 1	Interview 2	Level of Agreement	Reason
1	Somewhat negative	Neutral	Very positive	Very positive	Moderately positive	Neutral	Moderately Positive	Strongly Agree	Same trend
2	Very positive	Very positive	Moderately Positive	Neutral	Moderately negative	Very positive	Very Positive	Strongly Disagree	Opposite trend
3	Neutral	Neutral	Neutral	Somewhat negative	Neutral	Somewhat Negative	Neutral	Somewhat Agree	Similar trend
4	Somewhat positive	Moderately positive	Moderately positive	Neutral	Neutral	Neutral	Somewhat Positive	Strongly Disagree	Opposite trend
5	Neutral	Moderately positive	Very positive	Very positive	Very positive	Neutral	Very Positive	Strongly Agree	Same trend
6	Moderately positive	Moderately positive	Moderately positive	Moderately positive	Moderately positive	Neutral	Somewhat Positive	Somewhat Agree	Similar trend
7	Neutral	Moderately positive	Neutral	Somewhat negative	Neutral	Moderately Positive	Somewhat Positive	Strongly Disagree	Opposite trend
8	Moderately negative	Somewhat negative	Neutral	Moderately negative	Somewhat Positive	Somewhat Negative	Neutral	Somewhat Agree	Similar trend
9	Somewhat negative	Somewhat positive	Moderately Positive	Neutral	Somewhat Positive	Somewhat Negative	Moderately Positive	Strongly Agree	Same trend
10	Neutral	Very positive	Very positive	Very positive	Very positive	Somewhat Positive	Very Positive	Strongly Agree	Same trend
11	Neutral	Very positive	Somewhat positive	Very positive	Very positive	Moderately Positive	Very Positive	Somewhat Agree	Similar trend
12	Moderately negative	Somewhat positive	Somewhat positive	Moderately Positive	Moderately positive	Somewhat Negative	Somewhat Positive	Somewhat Agree	Similar trend
13	Moderately negative	Very positive	Somewhat positive	Very positive	Very positive	Neutral	Somewhat Positive	Somewhat Agree	Similar trend
14	Very negative	Neutral	Neutral	Moderately negative	Somewhat negative	Moderately Negative	Moderately Negative	Somewhat Agree	Similar trend
15	Neutral	Moderately negative	Moderately positive	Somewhat Positive	Somewhat negative	Neutral	Neutral	Somewhat Agree	Similar trend
16	Moderately negative	Neutral	Somewhat Positive	Neutral	Neutral	Moderately Negative	Neutral	Strongly Agree	Same trend
17	Neutral	Somewhat positive	Very positive	Very negative	Somewhat positive	Moderately Positive	Neutral	Strongly Disagree	Opposite trend
18	Somewhat negative	Somewhat positive	Somewhat positive	Moderately positive	Moderately positive	Somewhat Negative	Somewhat Positive	Strongly Agree	Same trend
19	Neutral	Neutral	Somewhat positive	Moderately positive	Very positive	Moderately Negative	Somewhat Positive	Somewhat Agree	Similar trend

Table 43: Interviews and Course Assessments

Participant	Interview 1	Interview 2	Exams	Level of Agreement	Reason
1	Neutral	Moderately Positive	97, 99, 93, 95	Strongly Agree	Strong scores support increase
2	Very positive	Very Positive	99, 68, 74, 58	Strongly Disagree	Declining scores do not support
3	Somewhat Negative	Neutral	78, 58, 91, 86	Somewhat Agree	Scores support increase
4	Neutral	Somewhat Positive	96.5, 96, 70, 93	Somewhat Agree	Scores support increase
5	Neutral	Very Positive	100, 96, 100, 96	Strongly Agree	Strong scores support increase
6	Neutral	Somewhat Positive	96, 81, 84, 95	Strongly Agree	Scores support increase
7	Moderately Positive	Somewhat Positive	87, 79, 85, 92	Somewhat Agree	Fairly consistent scores support
8	Somewhat Negative	Neutral	91, 91, 82, 96	Somewhat Agree	Scores support increase
9	Somewhat Negative	Moderately Positive	97, 79, 81, 91	Somewhat Agree	Scores support increase
10	Somewhat Positive	Very Positive	94.5, 94.5, 95, 95	Strongly Agree	Strong scores support increase
11	Moderately Positive	Very Positive	96, 98, 99, 95	Strongly Agree	Strong scores support increase
12	Somewhat Negative	Somewhat Positive	97, 94, 95, 93	Strongly Agree	Strong scores support increase
13	Neutral	Somewhat Positive	100, 95, 87, 93	Strongly Agree	Strong scores support increase
14	Moderately Negative	Moderately Negative	85, 81, 55, 62	Strongly Agree	Scores support consistency
15	Neutral	Neutral	77, 73, 32, 66	Somewhat Agree	Scores support consistency
16	Moderately Negative	Neutral	95, 80, 55, 87	Somewhat Agree	Scores support slight increase
17	Moderately Positive	Neutral	96, 63, 66, 73	Somewhat Agree	Declining score support decrease
18	Somewhat Negative	Somewhat Positive	82, 69, 86, 89	Strongly Agree	Scores support increase
19	Moderately Negative	Somewhat Positive	93, 91, 89, 85	Strongly Agree	Scores support increase

Table 44: Interviews and Instructor Perspective

Participant	Interview 1	Interview 2	Instructor Perspective (Week 5, Week 10, Week 16)			Level of Agreement	Reason
1	Neutral	Moderately Positive	Mostly Positive	Mostly Positive	Very Positive	Strongly Agree	Same trend
2	Very positive	Very Positive	Very Positive	Mixed – Positive	Mixed-Negative	Strongly Disagree	Opposite trend
3	Somewhat Negative	Neutral	Very Negative	Mostly Negative	Mostly Positive	Somewhat Agree	Similar trend
4	Neutral	Somewhat Positive	Mostly Positive	Mostly Positive	Mixed	Somewhat Agree	Similar trend
5	Neutral	Very Positive	Mixed-Positive	Mixed-Positive	Very Positive	Strongly Agree	Same trend
6	Neutral	Somewhat Positive	Mixed	Very Positive	Very Positive	Somewhat Agree	Similar trend
7	Moderately Positive	Somewhat Positive	Mostly Positive	Mixed	Mixed-Negative	Somewhat Disagree	Similar trend, higher interviews
8	Somewhat Negative	Neutral	Mixed	Mixed-Positive	Mixed-Negative	Somewhat Agree	Similar trend
9	Somewhat Negative	Moderately Positive	Mixed	Very Positive	Mostly Positive	Somewhat agree	Similar trend
10	Somewhat Positive	Very Positive	Mostly Positive	Mostly Positive	Very Positive	Strongly Agree	Same trend
11	Moderately Positive	Very Positive	Mostly Positive	Mostly Positive	Mostly Positive	Somewhat Agree	Similar trend
12	Somewhat Negative	Somewhat Positive	Mixed	Mixed	Mixed-Positive	Somewhat Agree	Similar trend
13	Neutral	Somewhat Positive	Mixed	Mixed-Positive	Mixed	Somewhat Agree	Similar trend
14	Moderately Negative	Moderately Negative	Very Negative	Mixed-Positive	Very Negative	Strongly Agree	Same trend
15	Neutral	Neutral	Mostly Positive	Mixed-Negative	Mostly Negative	Somewhat Agree	Similar trend
16	Moderately Negative	Neutral	Mixed	Mixed-Negative	Mixed-Negative	Somewhat Agree	Similar trend
17	Moderately Positive	Neutral	Mixed	Mostly Negative	Mostly Negative	Somewhat Agree	Similar trend
18	Somewhat Negative	Somewhat Positive	Mostly Negative	Mixed	Mostly Positive	Strongly Agree	Same trend
19	Moderately Negative	Somewhat Positive	Mixed	Mixed-Positive	Mixed-Positive	Somewhat Agree	Similar trend

Table 45: Journals and Course Assessments

Participant	Journal 1	Journal 2	Journal 3	Journal 4	Journal 5	Exams	Level of Agreement	Reason
1	Somewhat negative	Neutral	Very positive	Very positive	Moderately positive	97, 99, 93, 95	Strongly Agree	Scores support trend
2	Very positive	Very positive	Moderately Positive	Neutral	Moderately negative	99, 68, 74, 58	Strongly Agree	Scores support trend
3	Neutral	Neutral	Neutral	Somewhat negative	Neutral	78, 58, 91, 86	Somewhat Agree	Scores support trend
4	Somewhat positive	Moderately positive	Moderately positive	Neutral	Neutral	96.5, 96, 70, 93	Strongly Agree	Scores support trend
5	Neutral	Moderately positive	Very positive	Very positive	Very positive	100, 96, 100, 96	Strongly Agree	Scores support trend
6	Moderately positive	Moderately positive	Moderately positive	Moderately positive	Moderately positive	96, 81, 84, 95	Strongly Agree	Scores support trend
7	Neutral	Moderately positive	Neutral	Somewhat negative	Neutral	87, 79, 85, 92	Somewhat Agree	Scores support trend
8	Moderately negative	Somewhat negative	Neutral	Moderately negative	Somewhat Positive	91, 91, 82, 96	Somewhat Agree	Scores support trend
9	Somewhat negative	Somewhat positive	Moderately Positive	Neutral	Somewhat Positive	97, 79, 81, 91	Somewhat Agree	Scores support trend
10	Neutral	Very positive	Very positive	Very positive	Very positive	94.5, 94.5, 95, 95	Strongly Agree	Scores support trend
11	Neutral	Very positive	Somewhat positive	Very positive	Very positive	96, 98, 99, 95	Strongly Agree	Scores support trend
12	Moderately negative	Somewhat positive	Somewhat positive	Moderately Positive	Moderately positive	97, 94, 95, 93	Somewhat Agree	Scores support trend
13	Moderately negative	Very positive	Somewhat positive	Very positive	Very positive	100, 95, 87, 93	Strongly Agree	Scores support trend
14	Very negative	Neutral	Neutral	Moderately negative	Somewhat negative	85, 81, 55, 62	Strongly Agree	Scores support trend
15	Neutral	Moderately negative	Moderately positive	Somewhat Positive	Somewhat negative	77, 73, 32, 66	Somewhat Agree	Scores support trend
16	Moderately negative	Neutral	Somewhat Positive	Neutral	Neutral	95, 80, 55, 87	Somewhat Agree	Scores support trend
17	Neutral	Somewhat positive	Very positive	Very negative	Somewhat positive	96, 63, 66, 73	Somewhat Agree	Scores support trend
18	Somewhat negative	Somewhat positive	Somewhat positive	Moderately positive	Moderately positive	82, 69, 86, 89	Strongly Agree	Scores support trend
19	Neutral	Neutral	Somewhat positive	Moderately positive	Very positive	93, 91, 89, 85	Somewhat Agree	Scores support trend

Table 46: Journals and Instructor Perspective

Participant	Journal 1	Journal 2	Journal 3	Journal 4	Journal 5	Instructor Perspective (Week 5, Week 10, Week 16)			Level of Agreement	Reason (Week 5 is Journal 2)
1	Somewhat negative	Neutral	Very positive	Very positive	Moderately positive	Mostly Positive	Mostly Positive	Very Positive	Somewhat Agree	Same trend
2	Very positive	Very positive	Moderately Positive	Neutral	Moderately negative	Very Positive	Mixed – Positive	Mixed-Negative	Strongly Agree	Same trend
3	Neutral	Neutral	Neutral	Somewhat negative	Neutral	Very Negative	Mostly Negative	Mostly Positive	Somewhat Agree	Similar trend
4	Somewhat positive	Moderately positive	Moderately positive	Neutral	Neutral	Mostly Positive	Mostly Positive	Mixed	Strongly Agree	Same trend
5	Neutral	Moderately positive	Very positive	Very positive	Very positive	Mixed-Positive	Mixed-Positive	Very Positive	Somewhat Agree	Similar trend
6	Moderately positive	Moderately positive	Moderately positive	Moderately positive	Moderately positive	Mixed	Very Positive	Very Positive	Somewhat Agree	Similar trend
7	Neutral	Moderately positive	Neutral	Somewhat negative	Neutral	Mostly Positive	Mixed	Mixed-Negative	Strongly Agree	Same trend
8	Moderately negative	Somewhat negative	Neutral	Moderately negative	Somewhat Positive	Mixed	Mixed-Positive	Mixed-Negative	Somewhat Agree	Similar trend
9	Somewhat negative	Somewhat positive	Moderately Positive	Neutral	Somewhat Positive	Mixed	Very Positive	Mostly Positive	Somewhat Agree	Similar trend
10	Neutral	Very positive	Very positive	Very positive	Very positive	Mostly Positive	Mostly Positive	Very Positive	Strongly Agree	Same trend
11	Neutral	Very positive	Somewhat positive	Very positive	Very positive	Mostly Positive	Mostly Positive	Mostly Positive	Somewhat Agree	Similar trend
12	Moderately negative	Somewhat positive	Somewhat positive	Moderately Positive	Moderately positive	Mixed	Mixed	Mixed-Positive	Somewhat Agree	Similar trend
13	Moderately negative	Very positive	Somewhat positive	Very positive	Very positive	Mixed	Mixed-Positive	Mixed	Somewhat Agree	Similar trend in consistency on different scales
14	Very negative	Neutral	Neutral	Moderately negative	Somewhat negative	Very Negative	Mixed-Positive	Very Negative	Strongly Agree	Same trend
15	Neutral	Moderately negative	Moderately positive	Somewhat Positive	Somewhat negative	Mostly Positive	Mixed-Negative	Mostly Negative	Somewhat Agree	Similar trend

Table 46 (cont'd)

16	Moderately negative	Neutral	Somewhat Positive	Neutral	Neutral		Mixed	Mixed-Negative	Mixed-Negative		Somewhat Agree	Similar trend
17	Neutral	Somewhat positive	Very positive	Very negative	Somewhat positive		Mixed	Mostly Negative	Mostly Negative		Strongly Disagree	Similar trend
18	Somewhat negative	Somewhat positive	Somewhat positive	Moderately positive	Moderately positive		Mostly Negative	Mixed	Mostly Positive		Somewhat Agree	Similar trend
19	Neutral	Neutral	Somewhat positive	Moderately positive	Very positive		Mixed	Mixed-Positive	Mixed-Positive		Somewhat Agree	Similar trend

Table 47: Course Assessments and Instructor Perspective

Participant	Instructor Perspective (Week 5, Week 10, Week 16)			Exams	Level of Agreement	Reason (Week 5 is Exam 1)
1	Mostly Positive	Mostly Positive	Very Positive	97, 99, 93, 95	Strongly Agree	Scores support observations
2	Very Positive	Mixed – Positive	Mixed-Negative	99, 68, 74, 58	Strongly Agree	Scores support observations
3	Very Negative	Mostly Negative	Mostly Positive	78, 58, 91, 86	Strongly Agree	Scores support observations
4	Mostly Positive	Mostly Positive	Mixed	96.5, 96, 70, 93	Strongly Agree	Scores support observations
5	Mixed-Positive	Mixed-Positive	Very Positive	100, 96, 100, 96	Strongly Agree	Scores support observations
6	Mixed	Very Positive	Very Positive	96, 81, 84, 95	Strongly Agree	Scores support observations
7	Mostly Positive	Mixed	Mixed-Negative	87, 79, 85, 92	Somewhat Agree	Scores support observations
8	Mixed	Mixed-Positive	Mixed-Negative	91, 91, 82, 96	Somewhat Agree	Scores support observations
9	Mixed	Very Positive	Mostly Positive	97, 79, 81, 91	Somewhat Agree	Scores support observations
10	Mostly Positive	Mostly Positive	Very Positive	94.5, 94.5, 95, 95	Strongly Agree	Scores support observations
11	Mostly Positive	Mostly Positive	Mostly Positive	96, 98, 99, 95	Strongly Agree	Scores support observations
12	Mixed	Mixed	Mixed-Positive	97, 94, 95, 93	Strongly Agree	Scores support observations
13	Mixed	Mixed-Positive	Mixed	100, 95, 87, 93	Somewhat Agree	Scores support observations
14	Very Negative	Mixed-Positive	Very Negative	85, 81, 55, 62	Strongly Agree	Scores support observations
15	Mostly Positive	Mixed-Negative	Mostly Negative	77, 73, 32, 66	Strongly Agree	Scores support observations
16	Mixed	Mixed-Negative	Mixed-Negative	95, 80, 55, 87	Somewhat Agree	Scores support observations
17	Mixed	Mostly Negative	Mostly Negative	96, 63, 66, 73	Strongly Agree	Scores support observations
18	Mostly Negative	Mixed	Mostly Positive	82, 69, 86, 89	Strongly Agree	Scores support observations
19	Mixed	Mixed-Positive	Mixed-Positive	93, 91, 89, 85	Strongly Agree	Scores support observations

REFERENCES

REFERENCES

Alldredge, J. R., & Brown, G. R. (2006). Association of course performance with student beliefs: An analysis by gender and instructional software environment. *Statistics Education Research Journal, 5*(1), 64-77.

Alves, M., Rodrigues, C. S., Rocha, A. M. A., & Coutinho, C. (2016). Self-efficacy, mathematics' anxiety and perceived importance: an empirical study with Portuguese engineering students. *European Journal of Engineering Education*, *41*(1), 105-121.

Alward, T. S. (2012). *A phenomenological study of student engagement at a for-profit, online college* [Doctoral dissertation, Northcentral University]. ProQuest Dissertations Publishing.

Ates, H., & Saylan, A. (2015). Investigation of pre-service science teachers' academic self-efficacy and academic motivation toward Biology. *International Journal of Higher Education*, *4*(3), 90-103.

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

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

Benken, B. M., Ramirez, J., Li, X., & Wetendorf, S. (2015). Developmental mathematics success: Impact of students' knowledge and attitudes. *Journal of developmental education*, 14-31.

Betz, N. E., & Hackett, G. (1983). The relationship of mathematics self-efficacy expectations to the selection of science-based college majors. *Journal of Vocational behavior*, *23*(3), 329-345.

Betz, N. E., & Hackett, G. (1993). *Manual for the mathematics self-efficacy scale: Sampler set*. Mind Garden.

Bong, M., & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: How different are they really? *Educational psychology review*, *15*(1), 1-40.

Brewer, D. S. (2009). *The effects of online homework on achievement and self-efficacy of college algebra students* [Doctoral dissertation, Utah State University]. All Graduate Theses and Dissertations. 407.

Cafarella, B. (2016). Acceleration and compression in developmental mathematics: Faculty viewpoints. *Journal of Developmental Education*, *39*(2), 12.

Cardetti, F., Khamsemanan, N., & Orgnero, M. C. (2010). Insights regarding the usefulness of partial notes in mathematics courses. *Journal of the Scholarship of Teaching and Learning*, 80-92.

Christensen, W. M., & Thompson, J. R. (2012). Investigating graphical representations of slope and derivative without a physics context. *Physical Review Special Topics-Physics Education Research*, *8*(2), 023101.

Connors, R., Smith, T., & Hickman, C. (2004). The Oz principle. New York, NY: Penguin.

Cordero, E. D., Porter, S. H., Israel, T., & Brown, M. T. (2010). Math and science pursuits: A self-efficacy intervention comparison study. *Journal of Career Assessment*, *18*(4), 362-375.

Dennis, L. M. G. (2013). *Multiplication achievement and self-efficacy in third-and fifth-grade students: Effects of cross-age peer tutoring and skill training* [Doctoral dissertation, The University of Arizona]. The University of Arizona.

Deshler, J., & Fuller, E. (2016). The effects of migration to a blended self-paced format for a remedial pre-college algebra mathematics course. *Journal of Computers in Mathematics and Science Teaching*, *35*(2), 113-129.

Ernst, J. V., Bowen, B. D., & Williams, T. O. (2016). Freshman engineering students at-risk of non-matriculation: Self-efficacy for academic learning. *American Journal of Engineering Education*, *7*(1), 9-18.

Fong, C. J., Zientek, L. R., Ozel, Z. E. Y., & Phelps, J. M. (2015). Between and within ethnic differences in strategic learning: A study of developmental mathematics students. *Social Psychology of Education*, *18*(1), 55-74.

Guuml, A. Ğ. I. R. G. A. N. (2013). Distance education students' self-efficacy levels of mathematical literacy. *Educational Research and Reviews*, *8*(14), 1137-1143.

Hall, J. M., & Ponton, M. K. (2005). Mathematics self-efficacy of college freshman. *Journal of Developmental Education*, *28*(3), 26-32.

Hekimoglu, S., & Kittrell, E. (2010). Challenging students' beliefs about mathematics: The use of documentary to alter perceptions of efficacy. *Primus*, *20*(4), 299-331.

Higbee, J. L. (1993). Developmental versus remedial: More than semantics. *Research and Teaching in Developmental Education*, *9*(2), 99-107.

Hodges, C. B. (2005). *Self-efficacy, motivational email, and achievement in an asynchronous mathematics course* [Doctoral dissertation, Virginia Tech]. VTechWorks.

Hodges, C. (2008). Self-efficacy, motivational email, and achievement in an asynchronous math

course. *Journal of Computers in Mathematics and Science Teaching*, *27*(3), 265-285.

Hodges, C. B., & Kim, C. (2010). Email, self-regulation, self-efficacy, and achievement in a college online mathematics course. *Journal of Educational Computing Research*, *43*(2), 207-223.

Jo, L. H. (2012). *Asian American college students' mathematics success and the model minority stereotype* [Doctoral dissertation, Columbia University]. Academic Commons.

Joët, G., Usher, E. L., & Bressoux, P. (2011). Sources of self-efficacy: An investigation of elementary school students in France. *Journal of educational psychology*, *103*(3), 649.

Johnson, P., & O'Keeffe, L. (2016). The effect of a pre-university mathematics bridging course on adult learners' self-efficacy and retention rates in STEM subjects. *Irish Educational Studies*, *35*(3), 233-248.

Kapasa, J., Samuels, E. B., & Mulendema, P. J. (2015). Problems faced by students and lecturers in teaching and learning of approximation in primary colleges of education: A case study of 4 primary colleges on the Copperbelt Province-Zambia. *Journal of Education and Practice*, *6*(17), 137-146.

Katz, S. (2015). Enhancing self-efficacy of elementary school students to learn mathematics. *Journal of Curriculum and Teaching*, *4*(1), 42.

Kung, H. Y. (2009). Perception or confidence? Self-concept, self-efficacy and achievement in mathematics: A longitudinal study. *Policy Futures in Education*, *7*(4), 387-398.

Larson, L. M., Pesch, K. M., Surapaneni, S., Bonitz, V. S., Wu, T. F., & Werbel, J. D. (2015). Predicting graduation: The role of mathematics/science self-efficacy. *Journal of Career Assessment*, *23*(3), 399-409.

Linnenbrink-Garcia, L., Perez, T., Barger, M. M., Wormington, S. V., Godin, E., Snyder, K. E., ... & Schwartz-Bloom, R. (2018). Repairing the leaky pipeline: A motivationally supportive intervention to enhance persistence in undergraduate science pathways. *Contemporary educational psychology*, *53*, 181-195.

Lopez, F. G., & Lent, R. W. (1992). Sources of mathematics self-efficacy in high school students. *The Career Development Quarterly*, *41*(1), 3-12.

Love, A. G. (2012). The growth and current state of learning communities in higher education. *New Directions for teaching and learning*, *2012*(132), 5-18.

Nguyen, G. N. T. (2015). A case study of students' motivation in college algebra courses. *Community College Journal of Research and Practice*, *39*(8), 693-707.

Nicolescu, R. (2015). *The analysis of self-efficacy for students enrolled in a calculus I course*

at a community college [Doctoral dissertation, University of Oklahoma]. SHAREOK.

Niemela, P., Mikkolainen, V., & Vuorinen, J. (2018). Compute mindlessly. Not! Map consciously. *Universal Journal of Educational Research*, *6*(11), 2669-2678.

Noble, J., & Sawyer, R. (2013). A Study of the effectiveness of developmental courses for improving success in college. ACT Research Report Series, 2013 (1). *ACT, Inc.*

Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Advances in health sciences education*, *15*(5), 625-632.

Olani, A., Hoekstra, R., Harskamp, E., & Van Der Werf, G. (2011). Statistical reasoning ability, self-efficacy, and value beliefs in a university statistics course. *Electronic Journal of Research in Educational Psychology*, *9*(1), 49-72.

Pajares, F. (1996). Self-efficacy beliefs and mathematical problem-solving of gifted students. *Contemporary educational psychology*, *21*(4), 325-344.

Pajares, F., & Graham, L. (1999). Self-efficacy, motivation constructs, and mathematics performance of entering middle school students. *Contemporary educational psychology*, *24*(2), 124-139.

Pajares, F., & Miller, M. D. (1994). Role of self-efficacy and self-concept beliefs in mathematical problem solving: A path analysis. *Journal of educational psychology*, *86*(2), 193.

Panaoura, A., Gagatsis, A., Deliyianni, E., & Elia, I. (2009). The structure of students' beliefs about the use of representations and their performance on the learning of fractions. *Educational Psychology*, *29*(6), 713-728.

Perry, D. R., & Steck, A. K. (2015). Increasing student engagement, self-efficacy, and meta-cognitive self-regulation in the high school geometry classroom: Do iPads help? *Computers in the Schools*, *32*(2), 122-143.

Pike, G. R., Kuh, G. D., & Massa-McKinley, R. C. (2008). First-year students' employment, engagement, and academic achievement: Untangling the relationship between work and grades. *NASPA journal*, *45*(4), 560-582.

Prindle, C. M. (2014). *A Qualitative Case Study Exploring Self-Efficacy and Its Influence on Fourth Grade Mathematics Students* [Doctoral dissertation, Northcentral University]. ProQuest Dissertations Publishing.

Raelin, J. A., Bailey, M. B., Hamann, J., Pendleton, L. K., Reisberg, R., & Whitman, D. L. (2014). The gendered effect of cooperative education, contextual support, and self-efficacy on undergraduate retention. *Journal of Engineering Education*, *103*(4), 599-624.

Siegle, D., & McCoach, D. B. (2007). Increasing student mathematics self-efficacy through teacher training. *Journal of Advanced Academics*, *18*(2), 278-312.

Smith, B. (2010). *A comparison of the cognitive and personal development of college-level mathematics students and developmental students at a community college* [Doctoral dissertation, The University of Memphis]. ProQuest Dissertations Publishing.

Thiel, T., Peterman, S., & Brown, M. (2008). Addressing the crisis in college mathematics: Designing courses for student success. *Change: The Magazine of Higher Learning*, *40*(4), 44-49.

U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics. (2018). *College student employment.* Retrieved from https://nces.ed.gov/programs/coe/indicator_ssa.asp.

Usher, E. L., & Pajares, F. (2008). Sources of self-efficacy in school: Critical review of the literature and future directions. *Review of educational research*, *78*(4), 751-796.

Usher, E. L. (2015). Personal capability beliefs. *Handbook of educational psychology*, *3*, 146-159.

Valencia-Vallejo, N., López-Vargas, O., & Sanabria-Rodríguez, L. (2018). Effect of motivational scaffolding on e-learning environments: self-efficacy, learning achievement, and cognitive style. *Journal of Educators Online*, *15*(1), n1.

Watts, B. K. (2011). *Relationships of mathematics anxiety, mathematics self-efficacy and mathematics performance of adult basic education students* [Doctoral dissertation, Capella University]. ProQuest Dissertations Publishing.

Williams, L. C., Conley, K., Pavletic, H., & Weller, K. (2020). C4 scholar program: Promoting success through accountability for at-risk students. *Innovative Higher Education*, 1-15.

Yurt, E. (2014). The predictive power of self-efficacy sources for mathematics achievement. *Education & Science/Egitim ve Bilim*, *39*(176).

Zientek, L. R., Fong, C. J., & Phelps, J. M. (2019). Sources of self-efficacy of community college students enrolled in developmental mathematics. *Journal of Further and Higher Education*, *43*(2), 183-200.