Student Attitudes in an Innovative Active Learning Approach in Calculus

Adam J. Castillo, Pablo Durán, Edgar Fuller, Charity Watson,
Geoff Potvin & Laird H. Kramer

STEM Transformation Institute, Florida International University, Miami, FL

A study conducted by the Mathematical Association of America showed that calculus not only has significant effects on students' decision to pursue STEM fields, but also on their attitudes towards mathematics. Inspired by this large-scale study, the present study sought to deepen the current understanding of the impact of calculus on student attitudes towards mathematics. Results of an implementation of the Modeling Practices in Calculus (MPC) model, an innovative active learning in mathematics (ALM) approach, in Calculus I at a large, urban, research intensive (R1) institution are presented. Using a randomized-control trial research design, students were randomly assigned to either traditional, lecture-based classrooms, or MPC classrooms. The Attitudes Towards Mathematics Inventory (ATMI) was used to measure student attitudes at the beginning and end of the course and results were compared from both MPC and traditional sections. Overall, MPC sections showed improvement over traditional instruction by having less negative impact on student attitudes. The enjoyment and self-confidence ATMI subscales showed significant differences at course completion for both semesters, when controlling for pre-ATMI score and term. Furthermore, the MPC model had a positive impact on female students' self-confidence as opposed to male students, acting as a gender equalizer.

Keywords: active learning, calculus, student attitudes, randomized control trial, gender equalizer

Subject classification codes: 97C20, 97D40

This is an original manuscript of an article published by Taylor & Francis in International Journal of Mathematical Education in Science and Technology on 04 Feb 2022, available at: https://doi.org/10.1080/0020739X.2021.2023771.

Introduction

Calculus plays an important role in higher education. It is a foundational course required for almost every science, technology, engineering, or mathematics (STEM) major, as well as most pre-medicine programs, and thus acts as a transition point for students in almost every STEM degree program. Yet, for many students, calculus presents an overwhelming barrier with a significant effect on their decision to pursue STEM fields (Bressoud et al., 2015; Kogan & Laursen, 2014). Students' struggles to succeed in calculus courses are even acknowledged by instructors, who see these courses as discouraging, often fast-paced, and covering a wide range of topics (Bressoud et al., 2013; Bressoud et al., 2015; Sonnert et al., 2015).

Active learning approaches have recently been seen as an important strategy to address the difficulties STEM students have with calculus courses. Recent work has provided strong evidence that STEM undergraduate courses with components that involve active learning in mathematics (ALM) tend to improve student learning and provide more equitable outcomes (Freeman et al., 2014). Mathematics education researchers are now making a case for increased active learning presence in mathematics courses (Laursen & Rasmussen, 2019), noting its success in introductory courses (Rasmussen et al., 2019).

Studies on the effectiveness of ALM approaches, when compared to lectures, have mainly focused on students' achievement; most excluding the impact these courses might have on students' attitudes. Some studies support a positive correlation between active learning and student attitudes in science courses (Armbruster et al., 2009; Hammad et al., 2020; Yuretich et al., 2001); however, in calculus, the effect seems to be the opposite. As shown in one of the few studies on student attitudes in calculus, where Bressoud et al. (2013) found that college calculus instructors that included components of active learning in their courses (e.g., group work, student explanation of thinking, word problems, whole-class discussion) had actually a small negative impact on students' attitude.

Understanding this impact in depth is extremely important, since student attitudes are not only linked to their achievement (Evans, 2007; Sorge & Schau, 2002; Yee, 2010), but they have also proven to play a predominant role in shaping students' persistence in STEM programs (Bressoud et al. 2013, Maltese & Tai, 2011), and have also been related to issues of equity. A study examining the effect of attitudes by gender, for instance, recently found that students' lack of confidence in mathematics has been linked to attrition rates in STEM programs (Ellis et al., 2016). The authors showed that the odds of a female being discouraged from continuing in calculus is 1.5 times greater than that for a male student. Further efforts in understanding the impact of active learning in other attitudinal traits and demographics variables are needed.

Inspired by the call to understand the impact of ALM on students' experience in calculus, the results of the first-year implementation of an ALM approach in Calculus I were investigated, at a large, urban, research-intensive university. Student attitudes towards mathematics were measured at the beginning and end of Calculus I, using a valid and reliable survey instrument. This Calculus I course integrated the Modeling Practices in Calculus (MPC) model, which combines ambitious and good teaching approaches (Mesa et al., 2015; Sonnert et al., 2015), grounding various active learning strategies, with a focus on the inquiry practices of mathematicians. These strategies were also supported by enhanced facilitation of learning by incorporating Learning Assistants (Otero et al., 2010; Talbot et al., 2015) in the classroom.

The present study focused on the following two research questions: (1) To what extent does the MPC model, an innovative ALM approach, change student attitudes in Calculus I when compared to traditional instruction? (2) To what extent does the MPC model change student attitudes when compared to traditional instruction by certain key demographics, including gender, precalculus proficiency, and class-standing?

Before discussing the details of this study, a brief summary of relevant literature on the main aspects of student attitudes, active learning and the MPC model is presented.

Literature Review

Defining Attitudes Toward Mathematics

Research on student affect has been motivated by the belief that student attitude plays a significant role in the learning of mathematics (Neale, 1969). There is not a clear consensus on theory used to describe student affect, because research has given more attention to developing instruments to measure attitudes (Kulm, 1980; McLeod, 1992). As a construct, attitude, which originally referred to aspects of posture that expressed emotion, is now viewed as a mental orientation because of how this construct has been applied metaphorically to the attitude of mind (Ruffell et al., 1998; Stanton et al., 2021). This brings up a cause for confusion since the field of modern psychology differentiates between the cognitive, affective and behavioral aspects of the mind.

Research in the late 1980s and early 1990s often used a simple definition of attitude toward mathematics as a positive or negative emotional disposition toward mathematics (Haladyna et al., 1993; McLeod, 1992). More recent views started to acknowledge a more complex nature of the construct (Leder, 1985). Ma and Kishor (1997) defined attitude towards mathematics as an aggregated measure of one’s liking of mathematics, one’s tendency to engage in mathematics, a belief that one is good or bad at mathematics, a belief in the usefulness of mathematics, and a belief in the importance of mathematics. This view of attitudes towards mathematics recognizes that attitude as a multidimensional construct (Hart, 1989). Similarly, Tapia and Marsh (2004, 2005) suggest that attitude towards mathematics refers to one's feelings and emotions towards mathematics including enjoyment, motivation, self-confidence, and value. In agreement with Ma and Kishor (1997) and Tapia and Marsh (2004, 2005), attitudes towards mathematics are defined in this study as a multidimensional construct including the beliefs one has towards mathematics, by the emotions and feelings that one associates with mathematics, and by one's behavior related to mathematics (Ajzen, 1988; Hart, 1989; Triandis, 1971). This view of attitude towards mathematics takes into consideration the four factors identified by Tapia and Marsh (2004, 2005): enjoyment, motivation, self-confidence, and value. Given the difficulties of proposing an appropriate definition of attitudes towards mathematics for all research on student affect (Kuhn, 1980), this last definition assumes the role of a working definition viewed as a construct useful to the research question (Daskalogianni & Simpson, 2000).

Measuring Attitudes Toward Mathematics

Over the last 60 years, there has been a large interest in research on student affect in mathematics (Zan et al., 2006). This interest has been echoed by The National Council of Teachers of Mathematics (1989) as it reiterated the key role of research on student affect in its standards for curriculum and evaluation (Commission on Standards for School Mathematics, 1989).

The predominant role of student attitudes in mathematics education is also reflected on several literature reviews (Aiken, 1970, 1976; Kulm, 1980; Reyes, 1984). Much of their focus has been on students' overall attitudes toward mathematics, rather than trying to describe and analyze specific components of the affective domain. Leder

(1987) noted that there are many kinds of mathematics and various feelings (e.g., anxiety, enjoyment, motivation) associated with the affective domain.

To account for the underlying components of student attitudes, a myriad of instruments to assess student affect in a comprehensive way have been developed over the past several decades (Chamberlin, 2010; McLeod 1992, 1994), including the Mathematics Attitude Scale (Aiken, 1970), Mathematics Anxiety Rating Scale (Richardson & Suinn, 1972), Fennema-Sherman Mathematics Attitude Scale (Fennema & Sherman, 1976), and the Mathematics Attitude Inventory (Sandman, 1980).

However, Tapia and Marsh (2004) suggested that the previously developed instruments focused mainly on anxiety, competence, or enjoyment of subject matter, excluding other important factors associated with attitudes towards mathematics, such as motivation, self-confidence, and value towards learning mathematics. To include these neglected factors, Tapia and Marsh (2004) developed a 40-item Attitudes Toward Mathematics Inventory (ATMI). The ATMI is one of the latest instruments developed to measure student affect (Chamberlin, 2010), and has been refined by factor analyses, identifying four main domains of student attitudes toward mathematics: enjoyment, motivation, self-confidence, and value (Lim & Chapman, 2013; Tapia & Marsh, 2004).

Referential values for the 40-item ATMI scale and its subscales have been reported for a sample of 544 high school students (Tapia & Marsh, 2004). Results showed an overall ATMI mean of 3.43 (SD = 0.72). The self-confidence factor had a mean of 3.41 (SD = 0.88) with scores from these 15 items producing a Cronbach's alpha of 0.95. The value factor had a mean of 3.84 (SD = 0.67) with scores from 10 items producing a Cronbach's alpha of 0.89. The enjoyment factor contains had a mean of 3.19 (SD = 0.81) with scores from these 10 items producing a Cronbach's alpha of 0.89. Lastly, the motivation factor had a mean of 3.20 (SD = 0.99) with scores from five items producing a Cronbach's alpha of 0.88. These descriptive statistics and reported Cronbach's alpha provide baseline values to compare to postsecondary students.

To determine whether the four-factor model would hold for a population of college students, a confirmatory factor analysis was conducted by Tapia and Marsh (2002) using the model defined by Tapia (1996). Results from this analysis on 134 undergraduate mathematics students yielded good model fit statistics and high Cronbach alpha coefficients for each subscale, ranging from 0.87 to 0.96, indicating that the four-factor model holds with a high level of measure reliability. Among all the instruments mentioned to measure student attitudes towards mathematics, the ATMI has shown to have not only some of the highest psychometric properties, but also its comprehensive nature contributes to its validity.

Results of studies in college that have used the ATMI are limited; however, other instruments previously mentioned have been reported and found unexpected and concerning. One of the most striking results comes from a study on Calculus conducted by the Mathematical Association of America (MAA) and sponsored by the National Science Foundation. In this study, the experience of students going through a semester of Calculus I led to a strong decrease in confidence, enjoyment and desire to continue mathematics at the research universities and undergraduate colleges (Bressoud et al., 2013). Furthermore, students at research universities showed the greatest decreases even though they were found to be better prepared.

Although a noticeable selection bias was found due to the fact that students who were inclined to answer the survey at the end of the term were the students who had been doing well in the course or had not dropped the course, the results are still concerning. It is surprising that introductory calculus was so discouraging, particularly for students who do well in the course. There are several additional contributing factors such as class size and course pacing that needs to be considered, but the contradictory MAA study's results suggest that much still needs to be done to understand change in students' attitudes after one-semester of calculus (Bressoud et al., 2013; Bressoud et al., 2015).

Studies examining attitudinal changes by specific student demographics have also been carried out. First, as one of the clearest results, gender was found to have a significant impact on students' attitudes towards mathematics. A study from Tapia and Marsh (2000) examined the effects of gender and math achievement on student attitudes towards mathematics. Using the original sample of 544 high school students from Tapia and Marsh's (2004) study, a multivariate analysis of variance revealed an overall significant effect of gender on self-confidence and motivation, with males scoring higher than females on self-confidence and motivation. Another study (Ellis et al., 2016) revealed a sharp decrease in confidence and enjoyment at the end of an introductory calculus course among males and females, with the drop in confidence having a stronger impact on women. This drop in self-confidence in turn, impacted their success in calculus leading to females, on average, not persisting in calculus, and therefore, STEM.

Second, research on how attitudes towards mathematics are affected by students' mathematics preparation and more specifically precalculus proficiency is scarce; although studies have explored this relationship in other science fields (Tobias, 1994), finding a linear relationship between both variables. The relationship between student attitudes at the end of the course and initial mathematics preparation, however, has been studied indirectly. On the one hand, mathematics preparation has been found to be correlated to student achievement (Sonnert et al., 2020); and student achievement, on the other hand, has been found to be correlated to student attitudes at the end of the course (Evans, 2007). The direction of both of these relationships implies that students with low mathematics preparation, in general, are more likely to finish the course with low grades, and then their attitudes at the end of the semester can also be expected to decrease.

Third, in a metanalysis of 56 publications (1988-2014) in primary or secondary education, positive effects of innovative teaching strategies such as active learning were found to be weaker for older students (Savelsbergh et al., 2016) which might suggest that the impact of active learning on student attitude could slightly vary by class standing. Last, regarding students' interest in pursuing STEM careers, attitudes towards mathematics plays a vital role. Studies indicate that crucial to maintaining the desire for a STEM career is that students grow to show both high levels of engagement and interest in mathematics and science (Lent et al., 1994; Fouad et al., 2002). The impact of student attitudes toward mathematics on STEM career persistence might imply that students' choice to pursue non-STEM careers can in turn negatively influence their attitudes towards mathematics (Tai et al., 2006).

Overall, correlations found between student attitudes toward mathematics and the demographic variables previously discussed (gender, precalculus proficiency, class standing and STEM choice) suggest that the influence of active learning classrooms on student attitudes could include relevant interactions. However, the study of the role of these interactions in college calculus when examining the impact of active learning on students' attitudes has often been neglected in the literature.

Impact of Instruction on Student Attitudes Toward Mathematics

A recent metanalysis of over 56 publications (1988-2014) found that innovative teaching approaches have positive effects on student attitudes and achievement

(Savelsbergh et al. 2016). Although these results were conducted in primary and secondary school, included different effect sizes by domain (mathematics, chemistry, biology and physics), and suggested a predominant role of the strategies' novelty on student attitudes.

Research in college calculus is more limited and suggest a different influence of ALM on student attitudes. In the Characteristics of Successful Programs in College Calculus (CSPCC) project, Sonnert et al. (2015) found that faculty who employed 'good teaching' practices (e.g., allowed time to understand difficult ideas, help outside of class, clarity in presentation and answering questions) had the most positive impact on student attitudes, particularly with those students who began with a weaker initial attitude. Bressoud et al. (2013) found that 'good teaching' practices, more aligned to lecture-based instruction, are most strongly correlated with maintaining student confidence in and enjoyment of mathematics.

On the other hand, this work found that faculty who employed 'ambitious teaching' practices more aligned to ALM (e.g., group work, student explanation of thinking, word problems, whole-class discussion) had a small negative impact on student attitude but a positive effect on students who already showed a positive attitude toward mathematics. The effect of such practices can be detrimental without good instructor to student relationships (Bressoud et al., 2013). Sonnert et al. (2015) suggest more probing into the impact of 'ambitious teaching' on students' attitudes, particularly because teaching characteristics were found to interact with class size.

Considering the still limited research available in college calculus, the negative impact of classrooms more aligned to AL strategies, such as ambitious teaching, on student attitudes is yet not conclusive. Furthermore, these results seem to contradict other studies carried out in other STEM fields. In physics, a longer tradition of well-documented use of active learning strategies had led to overall improvement of students' attitudes towards physics (Beichner & Saul, 2003; Beichner et al., 2007). In biology, AL classrooms not only have improved student attitudes but have also increased course engagement and student satisfaction (Armbruster et al., 2009). The contradiction between the positive influence of active learning classrooms on student attitudes in other disciplines in STEM and the preliminary results found in college calculus classrooms suggest a strong need for further research to explore the existence of relevant unexplained variables.

Although the underlying mechanisms of attitudinal changes in AL instruction have not been fully understood yet, studies on student interactions suggest that these changes could be indirectly caused by the influence of AL on the interactions students have with instructors and/or learning facilitators, and those they have with their peers.

On the one hand, as instructors' predominant role in AL shifts from presenting material to guiding students' learning process, AL classrooms naturally create more opportunities to provide tailored and timely feedback to students than traditional classrooms. This type of feedback would enhance instructors' social congruence (Schmidt & Moust, 2000), as students are more likely to perceive them as supportive, interested, and caring. Instructors' social congruence, in turn, has been clearly linked to positive changes in students' ATM (Rawnsley & Fisher, 1998; Rotgans & Schmidt, 2011).

On the other hand, student attitudes might also be influenced by the quality of interactions students have with their peers in AL classrooms. As these classrooms are based on engaging students into working in groups, the focus of the learning process tends to shift from individualistic and competitive to more collaborative and/or cooperative in nature. This shift has been shown to positively influence college students' quality of interpersonal interactions, self-esteem, and social support, as well as student attitudes in general (Johnson et al., 1998; Prince, 2004; Springer et al., 1999). Furthermore, the collaborative interactions students have in AL classrooms might also create opportunities for them to frequently provide feedback to their peers, a role that has been previously linked to increases in self-confidence (Astin, 1993; Sax, 1994).

The contrast between the more competitive learning environment in traditional calculus courses and the more collaborative environment found in AL classrooms, also suggests an explanation to attitudinal changes in specific students' groups, especially regarding gender. Some studies, for instance, have linked gender gaps in self-confidence to competitive learning environments (Federičová et al., 2018; Sax, 1994). Research on the effect of competitive learning environments on other demographic variables such as year in college, career interests, and race/ethnicity is still limited, and further studies are needed.

To summarize, studies on the influence of classroom interactions on student attitudes suggest a possible mechanism for attitudinal changes in AL classrooms. The design of AL approaches could be providing instructors with more opportunities for better formative assessment and creating a more collaborative learning environment for students. These features have been proven to positively influence student attitudes, particularly students from certain groups, suggesting a plausible mechanism for attitudinal changes due to AL.

In addition to the influence of AL on classroom interactions, it is also important to notice that the impact that AL can have on students' grades can also be contributing to their attitudinal changes (Pascarella et al., 1987; Smart & Pascarella, 1986). Further research to clearly differentiate between the effect of classroom interactions, student grades, and other unaccounted causes of attitudinal changes in AL classrooms is, however, still needed.

Modeling Practices in Calculus (MPC) Model

The traditional Calculus classroom at the institution where this study was carried out looks like typical mathematics classrooms at research universities: large class sizes that are often lecture-based (Bressoud et al., 2015; Scott et al., 2015). These types of classrooms often include minimal "good" teaching practices as noted by Sonnert et al. (2015). While these classrooms may include additional support in the form of teaching assistants, duties carried out usually include leading discussion sections, working problems in class, holding office hours, and grading course assignments and exams (Kung, 2010). Moreover, these lecture-based classrooms often foster passive learning (Bransford et al., 2000) and often fail to build student confidence, interest, and motivation (Weimer, 2002). Given the shortcomings associated with traditional-lectured based classrooms, the institution where this study took place felt the need to implement active learning in calculus.

The Modeling Practices in Calculus (MPC) model is the innovative active learning in mathematics (ALM) approach that accompanied this study. The authors of the study built this model with the supporting curriculum, by intentionally incorporating well-established recommendations for mathematics and calculus instruction, including 'ambitious teaching' practices and strategies promoted by national mathematics societies and national reports (Bressoud, 2015; Rasmussen et al., 2019; Sonnert et al., 2015). Faculty, who followed the MPC model, participated in professional development that included a summer workshop to introduce them to the MPC model prior to teaching and weekly planning meetings throughout the semester to support the model adoption.

The MPC model integrates three core elements: cooperative learning, social metacognition, and a culturally appropriate learning environment. The MPC model fosters ubiquitous cooperative learning where students work together to maximize their own and each other's learning through shared learning goals (Johnson et al., 2007, 2014; Slavin, 1996). Every class, students are asked to join in groups to work on a set of learning activities that lead students in mathematical investigations in groups. These mathematical investigations develop core calculus ideas such as limits, rates of change, related rates, and accumulation. While engaging into these series of learning activities, students proceed through a concept development with questions and problems designed to lead them to essential insights and challenge their mathematical toolset. Furthermore, to foster further cooperation, the MPC model situates instructors and Learning Assistants, trained 'near peer' undergraduate classroom facilitators (Otero et al., 2010), as key agents in centering mathematical discussions in and between groups, as they constantly interact with groups and have multiple opportunities to redirect students' questions and comments to the group. Instructors are also encouraged to ask students to present ideas or solutions to problems in class to other groups using whiteboards, which also emphasizes student cooperation as they work collaboratively towards a common goal, strengthening their identity as part of a group. In such collaborative space centered on students, the time dedicated to lecturing is minimized to only general remarks made by the instructor that in most cases, accounts for less than ten percent of the class period.

For example, in one class, after students are led through group discussions to have an intuitive understanding of the basic ideas of average rate of change, instantaneous rate of change, and one- and two-sided limits, they are then asked to work in groups to develop an idea of how to compute limits without the need for a graphical or tabular representation. The learning activity then presents a question involving piecewise functions and asks students to compute one- and two-sided limits without any visual representations. By working together on the learning activity, students note how they use the domain of the piecewise function to 'visualize' what happens which helps them build an understanding of direct substitution, the first technique for computing limits presented.

The MPC model also includes social metacognition as an essential element, because research shows that this domain should not only include beliefs about one's own knowledge, emotions, and actions, but should consider the beliefs about others' knowledge, emotions, and actions (Jost et al., 1998). Benefits of social metacognition include distributing metacognitive demands among group members, making metacognition visible, and improving individual cognition (Chiu & Kuo, 2009). The MPC model promotes the development of social metacognition by supporting group members' identification of mistakes and construction of shared knowledge when working through learning activities. During these activities students are asked to reflect on the mathematical concepts of the day, describe these concepts to their peers in their own words, choose appropriate problem-solving strategies, validate peers' ideas within the groups and as a class. This last component is represented for instance when students in groups are asked to write up and present solutions they developed in their group on whiteboards, as group members must monitor each other's thinking and make suggestions to control their group problem solving.

The third core element of the MPC model is that this it is intentionally designed as a culturally appropriate learning model, as it allows students to try out their ideas in a low-stakes, safe environment, receive ongoing formative feedback from an instructional team, and participate in a community of learners. The instructor establishes a safe learning environment by messaging to students from the first day and regularly that making mistakes and asking questions are acceptable and a natural part of mathematics. The low-stakes environment is also enhanced as Learning Assistants (LAs), or undergraduate classroom facilitators, are integrated into the classroom to support learning with groups (Otero et al., 2010; Talbot et al., 2015). LAs are natural agents of this culturally appropriate model, as their demographics are that of the students, who provide insights and connections from the point of view of a recent participant in the course. LAs provide an essential component that establishes a strong connection to the ways in which students are constructing knowledge and LAs also help to mitigate ‘blind spots’ that experienced mathematicians bring into dialogues. This mitigation helps to increase the flow of ideas from the students to instructors, so that discussions are more strongly centered on students’ points of view.

Finally, in the MPC model, instructors are asked to meet weekly with LAs to learn about the perceptions LAs have about students, including the main difficulties students may have had over the past week in learning course materials or collaborating in their groups. This valuable feedback helps instructors to perceive when more clarity is needed or when to provide additional examples to better illustrate a particular concept. As LAs constantly scaffold students’ learning, clarify misconceptions, encourage collaboration, and provide valuable feedback to instructors, they are seen in the MPC model as a vital piece that holds the class together as a learning community (Emenike et al., 2020).

The MPC model is designed to motivate both cognitive and affective arenas to learn calculus. Preliminary analyses show great results in student learning, as a result, this study wants to find out if attitudes towards mathematics come along too.

Methods

This study focused on Calculus I instruction at a large, urban, research-intensive (R1) university. It initially included a randomized control trial experiment during the Spring 2019 and Fall 2019 terms to establish strong, reliable evidence (Czaja & Blair, 2005), assessing the impact of the MPC model on students' attitudes in calculus.

Participants

The sample in this study consisted of a total of 553 Calculus I students. In the Spring 2019 semester, a total of 168 of these students were randomly assigned to three control and three treatment sections. In the subsequent semester, the number of sections increased due to semester enrollment trends and the gradual MPC model implementation design. In the Fall 2019 semester, the total of students participating in this study expanded to a total of 385 students randomly assigned to six control and six treatment sections. Treatment sections adopted the MPC model, while control sections represented traditional, lectured-based classrooms. Instructors in control sections were not guided to use any particular instructional practices, though lecture was the accustomed practice. The syllabi provided to all sections in both control and treatment groups were consistent in terms of content coverage and course expectations. Other elements that were common across both control and treatment groups included: course pacing, weekly schedule, instructional supports, and a cumulative final exam. Courses were also provided with additional instructional support which included teaching/learning assistants.

Additionally, students' demographics data were reported by students to the university and collected at the time of course enrollment. A breakdown of the number of students by treatment group and the demographics of all participating students, can be seen in Table 1.

Measures and Procedure

Measures

Student attitudes were measured using the Attitudes Towards Mathematics Inventory (ATMI) (Tapia & Marsh, 2004). This survey is composed of 40 items measuring the four subscales described in the previous section of this study: enjoyment (10 items), motivation (5 items), self-confidence (15 items), and value (10 items). Eleven items of this survey were reversed-coded later for data analysis. This instrument was chosen due to its ability to capture the most relevant dimensions of students' attitudes, and documented reliability in college settings.

As an auxiliary measure, students were also asked to take the Precalculus Concept Assessment (PCA) survey, a reliable and validated precalculus proficiency instrument (Carlson et al., 2010). The PCA survey consists of 25 multiple-choice items and its purpose was twofold: as an additional demographic variable, and as covariance, when imputing missing data. As a demographic variable, PCA scores of less than 13 were associated with low precalculus proficiency, while PCA scores of 13 or higher were associated with high precalculus proficiency. This cut-off score was suggested by Carlson et al. (2010), and it was based on the correlation between the PCA scores and passing grades of 248 calculus college students.

Procedures

In order to obtain the final sample in this study, students enrolled in multiple, 80-seat (twice the normal size) sections of introductory calculus, chosen to fit their schedules as they normally would. Instructor names were invisible to students throughout this enrollment process. Two days prior to the beginning of each term, each of these 80-seat sections were then split into two 40-seat sections by assigning each student at random to one of either a treatment (MPC) or control (non-MPC lecture-based traditional instruction) section. After this random assignment was completed, students were still allowed to change sections prior to the enrollment deadline.

In the Spring semester, a total of 261 students were randomly assigned to ten sections within the study, with 130 students of these in five treatment sections and 131 students in five control sections. In the Fall 2019 Semester, a total of 533 students were randomly assigned to 16 sections within the study, with 271 students in eight treatment sections and 258 students in eight control sections.

Since only sections with matching schedules (same day/time teaching) were included in this study, the final sample included three sections per treatment in the Spring 2019, and six sections per treatment in the Fall 2019 semester. A group of students, no larger than 21% of the sample, were not part of the original random assignment. This group included students from other sections who decided to enroll in RCT sections after the split and prior to the enrollment deadline. Although, the number of students enrolled after the split were more or less evenly distributed in control and treatment sections, less than a dozen female students switched from control to treatment sections, and less than a dozen male students switched from treatment to control sections. These switches were not possible to control, however, a larger number of students (and treatment sections) in the Fall semester led to a more balanced distribution in gender that minimized this bias.

The number of students per treatment per semester in the final sample can be seen in Table 1. Furthermore, as explained in the subsequent section, confirmatory analysis on the main research questions of this study were carried out with the true RCT sample to measure the impact of student switching sections.

Students were asked to complete the ATMI and the PCA survey, at the beginning (first week of classes) and end of the semester (last two weeks of classes). Surveys were administered by the instructors, following a protocol that involved ensuring students their participation was not going to influence their grade in any way.

Missing data

The overall unweighted unit response rate for both semesters was 82.7% for pre-surveys, 66.8% for post surveys, and 64.8% for students submitting both surveys. A breakdown of survey response rates by semester is presented in Table 2.

A total of 82 (8.8%) out of 924 pre- and post-surveys collected were partially completed (not all but at least one item completed). Excluding blank surveys, item non-response rate for both treatment and control groups for each item was less than 3.4% for both pre- and post-surveys. When including blank pre-ATMI surveys later paired with completed or partially completed post-ATMI surveys, on the other hand, the percentage of missing values across the 40 pre-ATMI items of collected surveys ranged between 9% to 14%. However, this percentage was much higher for the post-ATMI items, ranging between 18% to 37%. This unbalance was mainly explained by high rates of student attrition, especially in the control section.

Due to differences in attrition rates between control and treatment sections, missing data was considered missing not completely at random (Rubin, 1976). Potential loss of statistical power and biased estimates due to this missingness were addressed using a multiple imputation (MI) algorithm (Kang, 2013). The extensively validated expectation-maximization with bootstrapping (EMB) MI algorithm AMELIA II was used to impute unit and item non-response (Honaker et al., 2011).

MI was carried out considering the pre-post design as a time series, using the pre-PCA survey results as a covariate. Since the percentage of missing data was less than 30%, using more than 30 iterations was considered to be appropriate (White et al., 2011). However, 100 iterations were used to improve analysis power (Graham et al., 2007). Although assumptions of normality were violated for each survey item, given the sample size in this study and the presence of slight deviations from normality, the EMB algorithm was assumed to be robust against these violations (Demirtas et al., 2008; Schafer & Olsen, 1998).

Unit non-response was further addressed by combining MI with a proximity score matching (PSM) algorithm (Granger et al., 2019). For this purpose, the nonparametric pre-processing algorithm MatchIt (Stuart et al., 2011) was used with nearest neighbor matching and a 0.2 standard deviation of caliper (Austin, 2011).

Data Analysis

All the analyses carried out in this study were obtained after applying the MI and PSM algorithms to the ATMI raw data. Means, standard errors and statistical tests were all determined following the standard rules to pool results from multiply imputed datasets (Rubin, 1976).

Multiple factorial ANCOVA's were conducted to understand the impact of the MPC model on student attitudes compared to traditional lecturing, when controlling for pre-test ATMI scores and term—the effect of having students from two different semesters. Since specific attitudinal traits analysis have been proven to be relevant when understanding the complexity of student attitudes, specific analysis ANCOVA's were conducted on each ATMI subscale.

ATMI scale and its subscales of motivation, enjoyment, value, and self-confidence responses were considered interval variables, given their Likert scale nature and large sample size (Pell, 2005; Carifio & Perla, 2007, 2008). This assumption followed previous studies (Asante, 2012; Karjanto, 2017; Primi et al., 2020; Wilson &

Grigorian, 2019) and was supported by two main theoretical positions. First, the existence of participants with no attitudes towards mathematics was considered not possible for the scales and subscales used in the study (i.e., no scale was considered to have a true zero). Each item in the ATMI survey was scored with a minimum of 1 (strongly disagree) and a maximum of 5 (strongly agree). A score of 3, corresponding to the “Neutral” response, was also considered a student attitude (and not the lack of it), following the definition of student attitude used in this study and other similar studies on this scale. When adding each item’s score, the overall ATMI scale and its subscales raw scores ranged accordingly. For instance, the ATMI scale raw score ranged between 40 and 200, and the enjoyment and value subscales raw scores ranged between 10 and 50. To facilitate interpretation, all scales were expressed as their items’ mean score, ranging then between 1 and 5.

Second, the assumption of equal distance between points was deemed to be reasonable for a couple of reasons. On the one hand, the 5-point Likert items used a balanced design with a midpoint. The choices presented in each item (Strongly Disagree, Disagree, Neutral, Agree, and Strongly Agree) in each scale were regarded to contribute equally to its quality. On the other hand, confirmatory analyses on the ATMI survey (Lim & Chapman, 2013; Ngurah & Lynch, 2013) suggest that the contribution of each item to each scale (overall, motivation, enjoyment, self-confidence, and value scale) is fairly homogeneous. Since the Likert items on each scale in this study were not examined individually, but as summated scales, this homogeneity prevented certain items from over or underrepresenting the scale. Therefore, 1-point differences in scores were assumed to be similar. It is worth noting that this last assumption is of course a matter of debate. Particularly, the midpoint choice of neutral response, in particular, limits the ability of the instrument of capturing more accurate data (Chyung et al., 2017). Also, there are small differences in loadings from item to item in confirmatory analyses that can be reduced by adjusting items' weights in each scale (León-Mantero et al., 2020). However, the robustness of the summated scales, given the sample size, the number of points in each item, and the number of items in each scale in this study (Carifio & Perla, 2007, 2008; Pell, 2005), contributed to efficiently control for these limitations.

Although the distribution of the ATMI scale and its subscales of motivation, enjoyment, value, and self-confidence responses slightly deviated from normality, sample size and ANCOVA tests design ensured robustness against this violation.

The covariate, pre-ATMI scores, was independent of the treatment effects and, assessed by visual inspection of a scatter plot, a linear relationship between pre-test and post-test ATMI scores for each treatment group was found. In addition to this visual inspection, there was no significant interaction between the covariate and the treatment effect, ensuring homogeneity of regression slopes. Homogeneity of the residual variances for both treatment groups was assumed, as Levene's test was not significant ($p > 0.05$). Two-way ANCOVA's were carried out to compare the main effects of treatment group on post-test ATMI scores, while controlling for pre-test ATMI scores and term.

Alternative approaches to compare post-test scores such as gain scores and post-test-only analyses were less appropriate, considering that this study was randomized (Huck & McLean, 1975). Since this study did not focus on outcome variable selection, nor on variable system structure, multivariate analysis before each individual ANCOVA was not carried out (Huberty & Morris, 1992).

Additionally, the effect of students' gender, class standing, precalculus proficiency, and STEM choice on student attitudes was also studied. The effect of other variables such as race/ethnicity was not considered in this analysis, due to limited sample sizes of certain groups. Multiple factorial ANCOVA's were conducted to compare the main effects of the treatment group and each of the four demographic variables on post-test ATMI scores, while controlling for pre-test ATMI scores and term. In all main effects examined in this study, partial eta squared was the measure of effect size used (Richardson, 2011).

Results

This study mainly focused on student attitudinal changes in an active learning calculus course when compared to a lecture-based course as measured by pre- and post-ATMI surveys. Key findings were organized first by the overall and subscales' ATMI scores, and then by the effect of demographic variables on these scores, including gender, precalculus proficiency as measured by the PCA scale, and class standing.

Overall and subscales attitudinal changes

A difference in overall ATMI post-test scores between treatment (M_{pooled}=3.42, SE_{pooled}=0.06) and control (M_{pooled}=3.25, SE_{pooled}=0.06) sections was found. After controlling for the pre-test scores, this difference was statistically significant [F(1, 1007.9349)=3.1350, p=0.07693] with a small effect size (PES=0.01100). Student attitudes in the Fall were, in general, higher than in the Spring semester, but changes in these attitudes followed a similar pattern across the academic year. This pattern can be seen in the breakdown by subscales in Figure 1.

As shown in Figure 1, student attitudes decreased in all subscales in the aggregated data, however, the MPC sections' negative impact was significantly lower than control sections. The subscales of enjoyment and self-confidence showed significant differences at course completion for both semesters, when controlling for pre-test and term.

None of the interactions between treatment and term was significant for the overall ATMI or any subscales scores [ATMI: $F(1, 2111.1924)= 0.4944$, $p=0.48203$; SC: $F(1, 2915.6307)= 0.2338$, $p=0.62872$; ENJ: $F(1,0.48203)=0.3107$, $p=0.57730$; MOT: $F(1, 4164.8816)=0.3520$, $p=0.55301$; VAL: $F(1,1619.3735)= 1.2081$, $p=0.27187$]. Such low interactions were also reflected in subsequent analysis for each of the main demographic analysis; however, they were not reported to highlight the main results of this study.

The highest difference in post-test scores was found in students' self-confidence between treatment ($M_{pooled}=3.35$, $SE_{pooled}=0.07$) and control ($M_{pooled}=3.12$, $SE_{pooled}=0.07$) sections with a small effect size (PES=0.019). Overall and subscale score gains with effect sizes for both semesters are shown in Table 3.

Gender

As shown in Figure 2, differences in ATMI post-test subscales means after controlling for pre-test scores between control and treatment sections were inconsistent across gender. Differences in all subscales for female students were significant, except for value. Furthermore, female students showed positive gains in self-confidence at the end of the semester with a small effect size (PES=0.032559). For male students, on the other hand, the only significant difference was found in self-confidence. Further details on each subscale can be found in Table 4.

The MPC model additionally acted as a gender equalizer. On the one hand, no significant differences in post-test scores between female and male students were found in the treatment sections (PES<0.008, in all subscales). On the other hand, significant differences in all subscales but the value subscale were found in the control sections, with small effect sizes of gender (PES>0.014). Gender effect sizes for each subscale for control and treatment sections are shown in Table 5.

Precalculus proficiency

As mentioned in the Methods section, participants were categorized into high and low precalculus proficiency groups according to their PCA score at the beginning of the semester. Scores of less than 13 were associated with low precalculus proficiency, while scores of 13 or higher were associated with high precalculus proficiency. Sample sizes of these groups can be found in Table 1. This cut off score was suggested by Carlson et al. (2010). Equal-frequency discretization of these scores into three categories (low, medium, high) yielded similar results. The analysis of mean post-test scores when controlling for pre-test scores shows that MPC sections had a higher impact on student attitudes than control sections for all students. Post-test scores means are shown in Figure 3.

As shown in Table 6, this impact was particularly high for students with high precalculus proficiency scores, showing small effect sizes in all subscales. For the group of students with low precalculus proficiency, on the other hand, self-confidence was the only subscale with significant differences between control and treatment. When scores were divided into three categories (low, medium, high), the two lowest and the highest categories yielded similar treatment effect sizes as the low and high Carlson et al.'s (2010) categories, respectively.

Class standing and STEM-intended majors

The majority of students in both sections, as shown in Table 1, were either freshman students (those who have earned fewer than 30 semester hours prior to taking the course) or sophomore students (those who have earned at least 30 semester hours but fewer than 60 semester hours prior to taking the course). The MPC model had a higher impact on sophomore than freshman students. None of the differences in overall and subscales scores were found significant for these students. On the other hand, for sophomore students the motivation subscale was the only non-significant difference. The overall score and the other subscales scores were found significant with small effect sizes ranging between 0.01 and 0.04. Further details can be seen in Table 7.

Most students enrolled in Calculus I, as shown in Table 1, were STEM students (STEM-intended majors). In general, differences in post-test scores in treatment sections followed a similar trend than control sections when controlled for pre-test scores. Students in the MPC model had similar post-test scores than the control group in the overall ATMI scales and its subscales for both STEM and non-STEM students. However, a couple of exceptions to this pattern were found in the subscales of self-confidence and enjoyment. For self-confidence, small effect sizes were found for both STEM and non-STEM students; for the enjoyment subscale, on the other hand, small effect sizes were only found in non-STEM students. Further details can be seen in Table 8.

Each previous analysis on the overall sample and for each main demographic were also conducted on true RCT sample, the sample that excluded students switching sections after the randomization process. Due to the extension of these analyses and the reduced sample size, their results were not presented here. However, it is worth noting that these results were found to be consistent with the main conclusions presented in this section.

Discussion

Student experiences in introductory calculus have a significant impact on their decision to pursue STEM majors and on their beliefs and attitudes towards mathematics (Bressoud et al., 2013; Maltese & Tai, 2011; Ellis et al., 2016). Institutions at all levels of higher education lose many students when calculus acts as a gateway course, restricting access to these career options, even for highly-motivated and well-prepared students. Active learning (AL) has recently become a successful strategy addressing this issue (Freeman et al., 2014). The impact of active learning approaches on student attitudes in calculus courses is not yet fully understood. Sonnert et al. (2015), in one of the few studies to measure attitudinal changes in calculus, found small negative gains in student attitudes in a large sample of colleges across the nation. Findings in this study, however, did not focus on clearly comparing lecture-based classrooms with a classroom based on active learning approaches.

Furthermore, studies that have investigated the link between gender (Ellis et al., 2016), mathematics prior preparation (Tobias, 1994; Evans, 2007), class standing (Savelsbergh et al., 2016), and STEM choice (Fouad et al., 2002; Tai et al., 2006) and student attitudes suggest that the influence of AL approaches on attitudes could interact with these demographics. This interaction has however not been fully understood in AL calculus classrooms in college.

The present study examined the relationship between AL and student attitudes in general and by these specific demographics capturing a large cross section of students' attitudes. A reliable and validated survey instrument was administered to randomly assigned student populations in treatment and control sections of a Calculus I course. Treatment sections used the Modeling Practices in Calculus (MPC) model, an innovative active-learning Calculus approach; control sections, on the other hand, were based on traditional, lecture-based classes previously used by the instructors. The ATMI survey, on student attitudes toward mathematics, and the PCA survey, on students' precalculus proficiency, were both administered to students at the beginning and end of the semester.

Impact of the MPC model on student attitudes in introductory calculus

We noticed a difference in the overall ATMI post-test scores after controlling for the pre-test scores and term between the control and the MPC sections. We found this difference to be statistically significant with a small effect size.

Student attitudes decreased in all ATMI subscales, however, the negative impact on students' attitudes in MPC sections was significantly smaller than the negative impact on students' attitudes found in traditional lecture-based sections. This finding is relevant, because it suggests a disruption in a tradition of larger negative gains in lecture-based classrooms on student attitudes. A disruption particularly clearer in the subscales of self-confidence and enjoyment, in which post-test scores in MPC sections differed by over three standard deviations over the non-MPC sections; including positive gains in the Spring semester for self-confidence.

Since similar collaborative learning classroom environments has previously proven to positively impact students' attitudes (Ifamuyiwa & Akinsola, 2008; Zakaria et al., 2010), we argue that the impact of the MPC model on student attitudes is a direct consequence of its active learning component which promotes multiple opportunities for students to engage collaboratively in different levels of mathematical discussions. Students are repeatedly encouraged to pose questions, describe concepts in their own words, elaborate and argue about these concepts, solve problems, and make mistakes in a safe and collaborative environment with timely feedback from instructors.

Furthermore, it is also possible that the MPC model could be positively influencing students' attitudes because of the changes it makes in the nature of classroom interactions. Instructors in MPC classrooms have more opportunities to provide better formative assessment and promote a more collaborative learning environment. Due to the improved quality of formative assessment, students could be more likely to perceive instructors' care and support, a perception that, in turn, has been proven to lead to attitudinal gains (Rawnsley & Fisher, 1998; Rotgans & Schmidt, 2011). Additionally, collaborative learning environments, as the one constantly promoted in MPC classrooms, has proven to positively influence student attitudes in general, when compared to the more competitive learning environments found in traditional calculus courses (Astin, 1993; Johnson et al., 1998; Prince, 2004; Sax, 1994; Springer et al., 1999). Other potential mechanisms of attitudinal changes, such as the AL classrooms' impact on achievement, and the extent to which they might contribute to these changes still need to be further examined.

Variables that explain the smaller impact of the MPC model on students' motivation and their views on the relevance of mathematics, when compared to self-confidence or motivation are not clear and they might be related to potential curriculum limitations. It is also unclear the extent to which the relationship between students' career goals and the traditional calculus content selection used by the MPC model might influence these attitudinal traits. Implementation of the MPC model and other renewed curriculum with a higher emphasis on mathematical applications (Durán & Marshall, 2019; Ganter & Barker, 2004; Ganter & Haver, 2011) that appeals more closely to these majors' career needs could have a higher impact on these attitudes.

Differences in student attitudes in calculus based on key demographics

One of the key findings, when studying the interaction of gender and attitudes, was the existence of positive gains in self-confidence by female students in both terms. Significant differences between male and female students were also found in the subscales of enjoyment, and motivation. Additionally, the MPC model was found to serve as a gender equalizer, since there were no differences in post-test scores for all ATMI subscales between females and males, while the control sections showed significant differences in enjoyment, motivation, and self-confidence with small effect sizes.

Previous research on the effect of active learning on student attitudes in mathematics courses provided supporting results regarding gender. Female's affective gains were statistically identical to those of males, while females in non-active learning in mathematics courses reported substantially lower affective gains than did their male peers (Laursen et al., 2011, 2014). Laursen et al. (2014) noted that females in active learning in mathematics courses on average increase their confidence in doing mathematics and desire to continue mathematics. The results in the present study corroborate that active learning courses offered a level playing field that involved learning experiences which benefit males and females equally.

Although the reasons why active learning classrooms might positively influence female students' attitudes have not been fully studied in the literature, preferences in female students to less individualistic and more collaborative learning environment could suggest a plausible explanation (Federičová et al., 2018; Sax, 1994). Control sections in this study were traditional calculus classrooms, where most of the time the instructor was dedicated to presenting materials and leaving little space for collaboration. MPC sections on the contrary were diametrically opposed to this model, where instructors dedicated little time to lecture and instead focused on engaging students in small groups to work on understanding the main topics of the class and on solving problems together as a group. It is possible that such collaborative learning space not only led to gains in all groups but particularly on female students who have typically been neglected in traditional calculus classrooms (Sax, 1994). Qualitative studies focused on describing classroom interaction in detail and how collaborative spaces are changing learning processes in female students and other specific demographics groups are needed to further understand the links found in our study.

Small effect sizes of the MPC model in student attitudes were also found, when controlling for students' precalculus proficiency. Students with both low and high PCA pre-test scores were similarly affected. Although Wu et al. (2018) suggest that student attitudinal changes and their initial academic background are not necessarily linked, the results of this study showed that the impact on attitudes was more pronounced for students with high precalculus proficiency. This discrepancy could be partially explained by the limitations in the understanding of certain foundational concepts students with low PCA might have, when participating in mathematical discussions. Students struggling with functions' visual representation in the plane, for instance, would be in a very different position than their peers to argue about derivatives' geometric representation. Such limitations might lead them to gradually decrease their presence in their groups. Sonnert et al. (2015) suggests this process slowly progresses throughout the semester as "the levels of challenge and expected performance are raised" (p. 384). Students with high precalculus proficiency, on the other hand, having a good understanding of these concepts are in a better position to engage in mathematical discussions. Furthermore, multiple opportunities for participation in these discussions throughout the semester, such as those found in active learning classrooms, would naturally lead to reinforcing their self-confidence.

The results on the variables of class standing and STEM major intent were not as conclusive as the results found in gender and precalculus proficiency, suggesting the presence of confounding variables. First, the impact on students' attitudes for both MPC and non-MPC sections was similar for freshman students. Conversely, MPC sections had a higher impact on student attitudes than non-MPC sections for sophomore students, but only for the subscales of self-confidence and enjoyment. It is possible that sophomore students' social adjustment and integration to college settings might have influenced their group interactions through the semester with peers and the instructors (Sharma, 2012). Second, the effect of the MPC model on STEM intended-majors was found to be similar to non-STEM intended-majors. However, the unbalanced sample size of both groups suggests the need of further research in order to draw any further conclusions.

Implications and Future Research

The results of this study have three main implications for education researchers in undergraduate mathematics. First, this study was quantitative in nature and did not provide further insight into reasons why students have attitudinal changes in both control and treatment sections. However, students in the treatment sections used the MPC model that is grounded in various active learning strategies with a focus on the inquiry practices of mathematicians.

A qualitative component, involving observations, which are critical for qualitative research (Marshall & Rossman, 2011), is needed to provide further insight on how students interacted with their peers and their instructors in both control and treatment sections. Qualitative studies that carefully examine classroom mathematical practices, and students' participation in mathematical activity, will contribute to highlighting the main reasons for these changes. Future research could also identify students from this study who saw both increases and decreases in ATMI overall and subscales to describe their course experiences and classroom interactions.

Second, the MPC model was designed to focus on instruction that is culturally relevant and to empower the classroom experience. Students in MPC sections were able to build their own understanding of calculus content where learning was more accessible, relevant, and meaningful for students. These outcomes are encouraging indicators of that experience for students. Future work will examine components associated with culturally relevant practices to determine their impact on students' experiences in calculus. Hagman (2021) also calls on mathematics departments to integrate practices related to diversity, equity, and inclusion to make their calculus programs more successful for students of color and women, as well as other populations of students underserved by their departments.

Last, Learning Assistants (LAs), who were used in MPC sections, are a core component of the innovative active learning in mathematics approach used in this study. LAs are natural agents of culturally relevant learning, providing insights and connections from the point of view of a recent participant in the course. LAs help students develop skills such as creating and defending ideas, making connections between concepts, and solving conceptual problems (Alzen et al., 2018). LAs provide an essential component that establishes a strong social connection surrounding the mathematical work being performed, and these connections enhance student ownership of the resultant mathematical knowledge. Future work should examine the impact of using LAs on students' attitudes in active learning in mathematics courses and the potential link between the use of LAs and changes in students' confidence and enjoyment as found in this study.

Limitations

Most limitations in this study are related to the difficulties found in trying to conduct a randomized controlled trial in college and implementing an innovative active learning approach. First, it is important to acknowledge the existence of two mathematics education researchers teaching one MPC section in each semester, since their specific training might have biased student attitudes (Andrews et al., 2011). The design of the MPC implementation, however, addressed this issue increasing the number of MPC instructors in the second semester from one to four instructors. Since the number of students in each section was similar for all sections, the effect of mathematics education researchers serving as MPC instructors did not seem to have a significant effect on student attitudes in the MPC sections overall results. Also, we did not see any significant variation in sections, so it is reasonable to conclude that all faculty implemented the model similarly.

Second, gaps in pre-test scores from treatment and control sections for the subscales of enjoyment and value were found in the Spring semester. The variables involved were not able to fully explain these gaps. Presumably, differences in instructor's messaging between treatment and control sections, as well as the much higher unit non-response of the control section post-ATMI might explain these gaps. Although the 30% of missing data found is a reasonable non-response, surveys were missing at a slightly higher rate for the control section over the first semester. This unbalance was mainly explained by a much higher dropout rate of control sections. The effect of missing data and pre-test scores' gaps found were both addressed using multiple imputation and proximity score matching. Only precalculus proficiency, as measured by PCA, was the covariance used in both algorithms; further studies should include other predictors of student attitudes including high-school GPA, SAT, and ACT.

Third, since students' demographics data was collected using the university's student information system, this study only considered binary gender descriptions and limited race and ethnicity sets of variables. This is a considerable limitation when addressing questions of equity (Chetkovich, 2019). Further studies that incorporate more comprehensive student demographic information will allow a better representation of the variety of groups not included in this study.

Fourth, this study was limited in its ability to capture what instruction looked like in both treatment and control sections. In terms of alignment between written and implemented curricula, this study did not include direct measures of fidelity of implementation. While we did not include measures of instructional fidelity, this limitation was partially addressed by professional development that included a summer workshop prior to teaching and weekly planning meetings throughout the semester to support a more reliable model adoption.

Last, although the study was initially designed as an RCT study, the ability of students to freely move to a different section limited the randomization of the study. However, subsequent analyses with the true RCT smaller sample were carried out on the main research questions to address this limitation, confirming the main results discussed here.

Conflict of Interest

No potential conflict of interests was reported by the authors.

References

Aiken, L. R. (1970). Attitudes toward mathematics. *Review of Educational Research, 40*(4), 551-596.

Aiken, L. R. (1976). Update on attitudes and other affective variables in learning mathematics. *Review of Educational Research, 46*(2), 293-311.

Ajzen, I. (1988). *Attitudes, personality, and behavior*. New York, NY: Open University Press.

Alzen, J., Langdon, L., & Otero, V. (2018). A logistic regression investigation of the relationship between the Learning Assistant model and failure rates in introductory STEM courses. *International Journal of STEM Education, 5*(56), 1-12.

Andrews, T. M., Leonard, M. J., Colgrove, C. A., & Kalinowski, S. T. (2011). Active Learning Not Associated with Student Learning in a Random Sample of College Biology Courses. *CBE—Life Sciences Education, 10*(4), 394-405.

Asante, K. O. (2012). Secondary students' attitudes towards mathematics. *IFE PsychologIA : An International Journal*, *20*(1), 121–133, https://doi.org/10.10520/EJC38916.

Astin, A. W. (1993). *What Matters in College? Four Critical Years Revisited.* San Francisco, CA: Jossey-Bass.

Armbruster, P., Patel, M., Johnson, E., & Weiss, M. (2009). Active Learning and Student-centered Pedagogy Improve Student Attitudes and Performance in Introductory Biology. *CBE—Life Sciences Education, 8*(3), 203-213.

Austin, P.C. (2011). Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. *Pharmaceutical Statistics, 10*(2), 150-161.

Beichner, R. J., & Saul, J. M. (2003). Introduction to the SCALE-UP (student-centered activities for large enrollment undergraduate programs) project. *Proceedings of the International School of Physics ''Enrico Fermi''.* Varenna, Italy. http://www. ncsu.edu/per/scaleup.html.

Beichner, R. J., Saul, J. M., Abbott, D. S., Morse, J. J., Deardorff, D., Allain, R. J., Bonham, S. W., Dancy, M. H., & Risley, J. S. (2007). The student-centered activities for large enrollment undergraduate programs (SCALE-UP) project. *Research-Based Reform of University Physics*, 1(1), 2–39.

Bransford, J.D., Brown, A.L., & Cocking, R.R. (2000). How People Learn: Brain, Mind, Experience, and School. Committee on Developments in the Science of Learning. Washington, DC: National Academies Press.

Bressoud, D. M., Carlson, M. P., Mesa, V., & Rasmussen, C. (2013). The calculus student: Insights from the Mathematical Association of America national study. *International Journal of Mathematical Education in Science and Technology*, *44*(5), 685-698.

Bressoud, D., Mesa, V., & Rasmussen, C. (2015) *Insights and recommendations from the MAA national study of college calculus.* MAA Notes. Washington, DC: The Mathematical Association of America.

Carifio, J., & Perla, R. (2007). Ten common misunderstandings, misconceptions, persistent myths and urban legends about Likert scales and Likert response formats and their antidotes. *Journal of Social Sciences*, *3*(3), 106-116.

Carifio, J., & Perla, R. (2008). Resolving the 50-year debate around using and misusing Likert scales. *Medical Education, 42*(12), 1150-1152.

Carlson, M., Oehrtman, M., & Engelke, N. (2010). The precalculus concept assessment: A tool for assessing students' reasoning abilities and understandings. *Cognition and Instruction, 28*(2), 113-145.

Chamberlin, S.A. (2010). A review of instruments created to assess affect in mathematics. *Journal of Mathematics Education, 3*(1), 167-182.

Chetkovich, C. (2019). How non-binary gender definitions confound (already complex) thinking about gender and public policy. *Journal of Public Affairs Education, 25*(2), 226-252.

Chiu, M. & Kuo, S. (2009). From metacognition to social metacognition: Similarities, differences, and learning. *Journal of Education Research, 3*(4), 1-19.

Chyung, S. Y., Roberts, K., Swanson, I., & Hankinson, A. (2017). Evidence-based survey design: The use of a midpoint on the Likert scale. *Performance Improvement, 56*(10), 15–23.

Commission on Standards for School Mathematics. (1989). *Curriculum and evaluation standards for school mathematics.* Reston, VA: National Council of Teachers of Mathematics.

Czaja, R. & Blair, J. (2005). Designing surveys: A guide to decisions and procedures (2nd Ed.). Thousand Oaks, CA: Sage.

Daskalogianni, K. & Simpson, A. (2000). Towards a definition of attitude: the relationship between the affective and the cognitive in pre-university students. In *Proceedings of the 24th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 217–224). Hiroshima, Japan.

Demirtas, H., Freels, S. A., & Yucel, R. M. (2008). Plausibility of multivariate normality assumption when multiply imputing non-Gaussian continuous outcomes: A simulation assessment. *Journal of Statistical Computation and Simulation, 78*(1), 69-84.

Durán, P. A., & Marshall, J. A. (2019). Mathematics for biological sciences undergraduates: a needs assessment. *International Journal of Mathematical Education in Science and Technology, 50*(6), 807-824.

Ellis, J., Fosdick, B., & Rasmussen, C. (2016). Women 1.5 times more likely to leave STEM pipeline after calculus compared to men: Lack of mathematical confidence a potential culprit. *PLoS ONE, 11*(7): e0157447.

Emenike, M., Schick, C.,Van Duzor, A., Sabella, M., Hendrickson, S., & Langdon, L. (2020). Leveraging undergraduate learning assistants to engage students during remote instruction: Strategies and lessons learned from four institutions. *Journal of Chemistry Education, 97*(9), 2502-2511.

Evans, B. (2007). Student attitudes, conceptions, and achievement in introductory undergraduate college statistics. *The Mathematics Educator, 17*(2), 24-30.

Federičová, M., Pertold, F., & Smith, M. L. (2018). Children left behind: Self-confidence of pupils in competitive environments. *Education Economics, 26*(2), 145–160.

Fennema, E. & Sherman, J. (1976). Fennema-Sherman mathematics attitudes scales: Instrument designed to measure attitudes toward mathematics, toward the learning of mathematics by females and males. *Journal for Research in Mathematics Education, 7*(5), 324-326.

Fouad, N., Smith, P., Zao, K. (2002) Across academic domains: Extensions of the social–cognitive career model. *Journal of Counseling Psychology*, *49*, 164–171.

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

Ganter, S. L., & Barker, W. (2004). A collective vision. The Curriculum Foundations Project: Voices of the Partner Disciplines, 1-10.

Ganter, S. L., & Haver, W. E. (2011). Partner discipline recommendations for introductory college mathematics and the implications for college algebra. Mathematical Association of America. http://www.maa.org/programs/faculty-and-departments/curriculum-department-guidelines-recommendations/crafty

Graham, J. W., Olchowski, A. E., & Gilreath, T. D. (2007). How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prevention Science, 8*(3), 206–213.

Granger, E., Sergeant, J. C., & Lunt, M. (2019). Avoiding pitfalls when combining multiple imputation and propensity scores. *Statistics in Medicine, 38*(26), 5120-5132.

Hagman, J.E., (2021). The eighth characteristics for successful calculus programs: Diversity, equity, and inclusion practices, *PRIMUS*, *31*(1), 70-90, https://doi.org/10.1080/10511970.2019.1629555.

Haladyna, T., Shaughnessy, J., Shaughnessy, M. (1983). A causal analysis of attitude toward Mathematics. *Journal for Research in Mathematics Education*, 14(1), 19-29.

Hammad, S., Graham, T., Dimitriadis, C., & Taylor, A. (2020). Effects of a successful mathematics classroom framework on students' mathematics self-efficacy, motivation, and achievement: A case study with freshmen students at a university foundation programme in Kuwait. *International Journal of Mathematical Education in Science and Technology*, https://doi.org/10.1080/0020739X.2020.1831091.

Hart, L. E. (1989). Describing the affective domain: Saying what we mean. In D. B. McLeod & V. M. Adams (Eds.), *Affect and mathematical problem solving* (pp. 37–45). New York, NY: Springer Verlag.

Honaker, J., King, G., & Blackwell, M. (2011). Amelia II: A program for missing data. *Journal of Statistical Software, 45*(7), 1-47.

Huberty, C. J., & Morris, J. D. (1992). Multivariate analysis versus multiple univariate analyses. In A. E. Kazdin (Ed.), *Methodological issues & strategies in clinical research* (pp. 351-365). American Psychological Association. https://doi.org/10.1037/10109-030.

Huck, S. W., & McLean, R. A. (1975). Using a repeated measures ANOVA to analyze the data from a pretest-posttest design: A potentially confusing task. *Psychological Bulletin, 82*(4), 511-518.

Ifamuyiwa, S. A., & Akinsola, M. K. (2008). Improving senior secondary school students' attitude towards mathematics through self and cooperative-instructional strategies. *International Journal of Mathematical Education in Science and Technology, 39*(5), 569-585.

Johnson, D., Johnson, R., & Smith, K. (2007). The state of cooperative learning in postsecondary and professional settings. *Educational Psychology Review, 19*(1), 15-29.

Johnson, D., Johnson, R., & Smith, K. (1998). Cooperative learning returns to college. What evidence is there that it works? *Change: The Magazine of Higher Learning, 30*(4), 26-35.

Johnson, D., Johnson, R., & Smith, K. (2014). Cooperative learning: Improving university instruction by basing practice on validated theory. *Journal of Excellence in College Teaching, 25*(3&4), 85-118.

Jost, J., Kruglanski, A., & Nelson, T. (1998). Social metacognition: An expansionist review. *Personality and Social Psychology Review, 2*(2), 137-154.

Kang, H. (2013). The prevention and handling of the missing data. *Korean Journal of Anesthesiology, 64*(5), 402-406.

Karjanto, N. (2017). Attitude toward mathematics among the students at Nazarbayev University Foundation Year Programme. *International Journal of Mathematical Education in Science and Technology*, *48*(6), 849–863, https://doi.org/10.1080/0020739X.2017.1285060.

Kogan, M., & Laursen, S. L. (2014). Assessing long-term effects of inquiry-based learning: A case study from college mathematics. *Innovative Higher Education*, *39*(3), 183-199.

Kulm, G. (1980). Research on mathematics attitude. In R. J. Shumway (Ed.), *Research in mathematics education* (pp. 356-387). Reston, VA: National Council of Teachers of Mathematics.

Kung, D. (2010). Teaching assistants learning how students think. In F. Hitt, D. Holton, & P. Thompson (Eds.), *Research in collegiate mathematics education VII* (pp. 143-169). Washington, DC.

Laursen, S., Hassi, M.L., Kogan, M., Hunter, A.B., & Weston, T. (2011). *Evaluation of the IBL mathematics project: Student and instructor outcomes of inquiry-based learning in college mathematics* (Report to the Educational Advancement Foundation and the IBL Mathematics Centers). Boulder, CO: Ethnography & Evaluation Research, University of Colorado Boulder.

Laursen, S. L., Hassi, M., Kogan, M., Weston, T. (2014). Benefits for women and men of inquiry-based learning in college mathematics: A multi-institutional study. *Journal for Research in Mathematics Education, 45*(4), 406-418.

Laursen, S. L., & Rasmussen, C. (2019). I on the prize: Inquiry approaches in undergraduate mathematics. *International Journal of Research in Undergraduate Mathematics Education*, *5*(1), 129-146.

Leder, G. (1985). Measurement of attitudes to mathematics. *For the Learning of Mathematics*, *34*(5), 18–21.

Leder, G. C. (1987). Attitudes towards mathematics. *The Monitoring of School Mathematics*, 2, 261–277.

Lent, R., Brown, S., & Hackett, G. (1994). Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *Journal of Vocational Behavior*, *45*, 79-122.

León-Mantero, C., Casas-Rosal, J. C., Pedrosa-Jesús, C., & Maz-Machado, A. (2020). Measuring attitude towards mathematics using Likert scale surveys: The weighted average. *PLOS ONE, 15*(10), e0239626, https://doi.org/10.1371/journal.pone.0239626.

Lim, S. Y., & Chapman, E. (2013). Development of a short form of the attitudes toward mathematics inventory. *Educational Studies in Mathematics*, *82*(1), 145-164.

Ma, X. & Kishor, N. (1997). Assessing the relationship between attitude toward mathematics and achievement in mathematics: A meta-analysis. *Journal for Research in Mathematics Education, 28*(1), 26-47.

Maltese, A. V., & Tai, R. H. (2011). Pipeline persistence: Examining the association of educational experiences with earned degrees in STEM among US students. *Science Education, 95*(5), 877-907.

Marshall, C., & Rossman, G. B. (2011). *Designing qualitative research* (5th ed.). Thousand Oaks, CA: Sage

McLeod, D. B. (1992). Research on affect in mathematics education: A reconceptualization. In D. Grows (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 575-596). New York, NY: McMillan.

McLeod, D. B. (1994). Research on affect and mathematics learning in the JRME: 1970 to the present. *Journal for Research in Mathematics Education, 25*(6), 637-647.

Mesa, V., Burn, H., & White, N. J. (2015). Good teaching of calculus I. In D. M. Bressoud, V. Mesa, & C. L. Rasmussen (Eds.), *Insights and recommendations from the MAA National Study of College Calculus* (pp. 83-91). Washington, DC.

National Council of Teachers of Mathematics. (1989). Curriculum and evaluation standards for school mathematics. Reston, VA: National Council of Teachers of Mathematics Commission on Standards for School Mathematics.

Neale, D. (1969). The role of attitudes in learning mathematics. *The Arithmetic Teacher*, *16*(8), 631–640.

Ngurah, A., & Lynch, D. P. (2013). A confirmatory factor analysis of attitudes toward mathematics inventory (ATMI). *The Mathematics Educator, 15*(1), 121–135.

Otero, V., Pollock, S., & Finkelstein, N. (2010). A physics department's role in preparing physics teachers: The Colorado learning assistant model. *American Journal of Physics, 78*(11), 1218-1224.

Pascarella, E. T., Smart, J. C., & Ethington, C. A. (1986). Long-term persistence of two-year college students. *Research in Higher Education, 24*(1), 47-71.

Pascarella, E. T., Smart, J. C., Ethington, C. A., & Nettles, M. T. (1987). The influence of college on self-concept: A consideration of race and gender differences. *American Educational Research Journal, 24*(1), 49-77.

Pell, G. (2005). Use and misuse of Likert scales. *Medical Education, 39*(9), 970–970.

Primi, C., Bacherini, A., Beccari, C., & Donati, M. A. (2020). Assessing math attitude through the Attitude Toward Mathematics Inventory – Short form in introductory statistics course students. *Studies in Educational Evaluation, 64*, 100838, https://doi.org/10.1016/j.stueduc.2020.100838.

Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education, 93*(3), 223–231.

Rasmussen, C., Apkarian, N., Hagman, J. E., Johnson, E., Larsen, S., & Bressoud, D. (2019). Characteristics of Precalculus Through Calculus 2 programs: Insights from a national census survey. *Journal for Research in Mathematics Education*, *50*(1), 98-112.

Rawnsley, D., & Fisher, D. L. (1998). Learning environments in mathematics classrooms and their associations with students' attitudes and learning. In *Annual Conference of the Australian Association for Research in Education*, Adelaide, Australia.

Rotgans, J. I., & Schmidt, H. G. (2011). The role of teachers in facilitating situational interest in an active-learning classroom. *Teaching and Teacher Education, 27*(1), 37–42, https://doi.org/10.1016/j.tate.2010.06.025.

Reyes, L. H. (1984). Affective variables and mathematics education. *The Elementary School Journal, 84*(5), 558-581.

Richardson, F. C. & Suinn, R. M. (1972). The Mathematics Anxiety Rating Scale: Psychometric data. *Journal of Counseling Psychology*, *19*(6), 551-554.

Richardson, J. T. (2011). Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review, 6*(2), 135-147.

Rubin, D. B. (1976). Inference and missing data. *Biometrika, 63*(3), 581-592.

Ruffell, M., Mason, J., & Allen, B. (1998). Studying attitude to mathematics. *Educational Studies in Mathematics*, *35*(1), 1–18.

Sandman, R. S. (1980). The mathematics attitude inventory: Instrument and user's manual. *Journal for Research in Mathematics Education*, *11*(2), 148-149.

Savelsbergh, E., Prins, G., Rietbergen, C., Fechner, S., Vaessen, B., Draijer, J., & Bakker, A. (2016). Effects of innovative science and mathematics teaching on student attitudes and achievement: A meta-analytic study, *Educational Research Review, 19*, 158-172,

Sax, L. J. (1994). Mathematical self-concept: How college reinforces the gender gap. *Research in Higher Education, 35*(2), 141–166.

Schafer, J. L., & Olsen, M. K. (1998). Multiple imputation for multivariate missing-data problems: A data analyst's perspective. *Multivariate Behavioral Research, 33*(4), 545-571.

Schmidt, H. G., & Moust, J. H. C. (1995). What makes a tutor effective - A structural-equations modeling approach to learning in problem-based curricula. *Academic Medicine, 70*(8), 708–714.

Scott, C., Green, L, & Etheridge, D. (2015). A comparison between flipped and lecture-based instruction in the calculus classroom. *Journal of Applied Research in Higher Education*, *8*(2), 252-264.

Sharma, B. (2012). Adjustment and emotional maturity among first year college students. *Pakistan Journal of Social and Clinical Psychology, 9*(3), 32-37.

Slavin, R. E. (1996). Research on cooperative learning and achievement: What we know, what we need to know. *Contemporary Educational Psychology, 21*(1), 43-69.

Sonnert, G., Sadler, P., Sadler, S., and Bressoud, D. (2015) The impact of instructor pedagogy on college calculus students' attitude toward mathematics. *International Journal of Mathematics Education in Science and Technology, 46*(3), 370-387.

Sorge, C., & Schau, C. (2002). *Impact of engineering students' attitudes on achievement in statistics: A structural model.* Presented at the Annual Meeting of the American Educational Research Association. New Orleans, LA.

Springer, L., Stanne, M. E., & Donovan, S. S. (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis. *Review of Educational Research, 69*(1), 21–51.

Stanton, J. D., Sebesta, A. J., & Dunlosky, J. (2021). Fostering metacognition to support student learning and performance. *CBE—Life Sciences Education*, *20*(2), 1-7.

Stuart, E. A., King, G., Imai, K., & Ho, D. (2011). MatchIt: nonparametric preprocessing for parametric causal inference. J*ournal of Statistical Software, 42*(8), 1-28.

Tai, R., Liu, C., Maltese, A., Fan, X. (2006). Planning early for careers in science. *Science, 312*, 1143–1144.

Talbot, R., Hartley, L., Marzetta, K., & Wee, B. (2015). Transforming undergraduate science education with learning assistants: Student satisfaction in large enrollment courses. *Journal of College Science Teaching, 44*(5), 28-34.

Tapia, M. (1996). *The Attitudes Toward Mathematics Instrument.* Presented at the Annual Meeting of the Mid-South Educational Research Association. Tuscaloosa, AL.

Tapia, M., & Marsh, G. (2000). *Effects of gender, achievement in mathematics, and ethnicity on attitudes towards mathematics.* Paper presented at the Annual Meeting of the Mid-South Educational Research Association, Bowling Green, KY. Retrieved from https://files.eric.ed.gov/fulltext/ED449044.pdf.

Tapia, M., & Marsh, G. (2002). *Confirmatory factor analysis of the attitudes toward mathematics inventory.* Paper presented at the Annual Meeting of the Mid-South

Educational Research Association, Chattanooga, TN. Retrieved from https://files.eric.ed.gov/fulltext/ED471301.pdf.

Tapia, M., & Marsh, G. (2004). An instrument to measure mathematics attitudes. *Academic Exchange Quarterly*, *8*(2), 16-22.

Tapia, M. & Marsh, G. (2005). Attitudes towards mathematics inventory redux. *Academic Exchange Quarterly, 9*(3), 272-275.

Triandis, H. C. (1971). *Attitudes and attitude change*. New York, NY: Wiley.

Weimer, M. (2002). Learner-Centered Teaching: Five Key Changes to Practice, San Francisco, CA: Jossey-Bass.

White, I. R., Royston, P., & Wood, A. M. (2011). Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine, 30*(4), 377-399.

Wilson, A. T., & Grigorian, S. (2019). The near-peer mathematical mentoring cycle: Studying the impact of outreach on high school students' attitudes toward mathematics. *International Journal of Mathematical Education in Science and Technology, 50*(1), 46–64, https://doi.org/10.1080/0020739X.2018.1467508.

Wu, X., Deshler, J., & Fuller, E. (2018). The effects of different versions of a gateway STEM course on student attitudes and beliefs. *International Journal of STEM Education, 5*(1): 44. https://doi.org/10.1186/s40594-018-0141-4

Yee, L. S. (2010). *Mathematics attitudes and achievement of junior college students in Singapore*. Paper presented at the Annual Meeting of the Mathematics Education Research Group of Australasia. Freemantle, Western Australia. Retrieved from https://files.eric.ed.gov/fulltext/ED521019.pdf

Yuretich, R. F., Khan, S. A., Leckie, R. M., & Clement, J. J. (2001). Active-learning methods to improve student performance and scientific interest in a large introductory oceanography course. *Journal of Geoscience Education, 49*(2), 111-119.

Zakaria, E., Chin, L. C., & Daud, M. Y. (2010). The effects of cooperative learning on students' mathematics achievement and attitude towards mathematics. *Journal of Social Sciences, 6*(2), 272-275.

Zan, R., Brown, L., Evans, J., & Hannula, M. (2006). Affect in mathematics education: An introduction. *Educational Studies in Mathematics, 63*(2), 113-121.

Figures

Tables