The Effects of After-School Peer Tutoring Programs in Mathematics

A THESIS
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Anna Jeanne Schneck

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF ARTS

Lesa Covington Clarkson

May 2010

Table of Contents

List of Tables

The Effects of After-School Peer Tutoring Programs in Mathematics

While educational paradigms shift, influenced by Federal and State administrations, field-tested research, and parental input, one question will always remain: how do we increase learning in mathematics for students? In particular, that question is more often referenced with those students that are traditionally underserved populations, such as non-White students; English language learners; students who require special education services; and students who have been tested for special education services but have not qualified, typically due to low IQs and low performance, with not enough discrepancy between the two. These are students who may have a negative perception of mathematics due to struggles and failed attempts at success. Regardless of their current or potential skill level, these students may view themselves as struggling with mathematics, especially when comparing themselves to their peers. However, a negative attitude can affect motivation, self-esteem, and future course and career choices.

For years, classroom educators have been experimenting with new ways to increase students' achievement in and attitudes about mathematics, such as computer games (O'Neal, Ernest, McLean, & Templeton, 1988) or creating small groups within larger classes (Aidman, 1997). Some methods are based on research; some are conceptualized intuitively by professionals in the field. Other strategies are handed down by experienced teacher mentors or suggested by colleagues. While many of these strategies for improving student achievement and attitude in mathematics have been researched individually, these strategies are often combined in a practical school setting. Intuitively, one may conclude that if individual methods have shown to be effective then the combination of one or more of these would result in similar, if not more positive, gains.

Practices that are designed to increase mathematics achievement and attitude in underserved populations have been investigated in isolated cases, most yielding favorable results (Baker, Rieg, & Clendaniel, 2006; Fuchs et al., 2005; O'Neal et al., 1988; Steele & Arth, 1998; Topping et al., 2003). More realistically, though, several interventions are combined if resources allow. Most schools want results, as measured in part by class performance, teacher analysis, and standardized test scores, and will utilize as many feasible strategies as possible. However, the average school tends to focus more on concrete, measurable increases in student learning, perhaps without much reflection on what specifically contributed to those increases.

The typical school is not concerned with *proving* the effectiveness of the implemented strategies. Therefore, research that focuses on the combination of interventions would be beneficial. This study analyzed the effects of a combination of interventions—peer tutoring and increased learning time—in an attempt to determine whether these two combined interventions would translate to growth in student achievement and a positive change in students' attitudes towards mathematics.

CHAPTER 1: Review of Literature

Peer Tutoring

Peer tutoring is a strategy that many schools implement within and outside the classroom. In their 2005 study, Dufrene, Noell, and Gilbertson define peer tutoring as "a student mediated instructional procedure in which student dyads or small learning groups work together on learning tasks (2).” Most research in peer tutoring focuses on the effect on either self-esteem or achievement (Baker et al., 2006; Fuchs et al., 2005; Topping et al., 2003; Walker, 2007), but none of these studies collect data on both self-esteem and academic achievement for the tutored students and a comparative control group.

The research that investigates the impact tutoring has on achievement reveals mixed results (Baker et al., 2006; Fuchs et al, 2005; Walker, 2007; Zuelke & Nelson 2001). The instruments and methods used to collect data varies; increases in mathematical achievement are often measured quantitatively through weekly assessments, course grades, grade point averages, pre- and posttests, or standardized test gains (Dufrene, Noelle & Gilbertson, 2005; Fuchs et al., 2005; Zuelke & Nelson, 2001). In most of these studies, the results have been largely positive. Fuchs, et al. (2005) conducted a study examining the effects of tutoring, though the tutors were not fellow students. Instead, they were college graduates, some with post-baccalaureate degrees. The researchers sorted students into three groups: Not at risk (NAR), At-risk control (AR control), and At-risk tutored (AR tutored). The researchers found that the NAR group started with the highest achievement levels on the pretest and maintained the highest performance on the posttest. The comparison between the two AR groups yielded promising results, however. The AR control group performed comparably on the pretest as the AR tutored group (mean scores within 0.02 of one another, with the AR control being slightly lower), but the AR tutored group showed more improvement as measured by the posttest by surpassing the AR control mean score by 2.4 points with an effect size of 0.33. Similarly, Baker, Reig, and Clendaniel (2006) reported positive results in their

20-week tutoring study, pairing college students with elementary students. The regular classroom teachers who had the tutees in class informally reported through interviews positive gains in both academic achievement and attitude. In addition, the results of the formal assessments were positive; the students demonstrated growth and improvement in both the math inventory and the self-esteem assessments. Unfortunately, there was no distinction in either formal assessment between students who received the tutoring services and those who did not, leaving no way to discern whether the results for the tutored students surpassed the results of those students who did not receive services.

However, not all instances of tutoring suggest that it contributes to an increase in mathematics achievement. Most notably, Zuelke and Nelson (2001) conducted a four-year study that analyzed the effects of community-based tutoring programs and found no gains in mathematical achievement during any year of the four-year study; in fact, in some years the math grade point averages of the participants were lower. In addition, the researchers found that there was a negative correlation between the number of hours tutored and math achievement gain. Several factors may be responsible for such unimpressive results, which will be discussed later in this paper.

While some studies focus on the effects of tutoring on achievement, others focus on self-esteem regarding mathematics. Gains in self-esteem may be measured a number of ways, one of which is through qualitative methods such as open-ended questionnaires or interviews (Baker et al., 2006; Walker, 2007). The results of these studies have mainly shown tutoring to have a positive effect on self-esteem for both the tutor and the tutee. Some studies utilize a quantitative approach, either in conjunction with (Baker, et al.,

2006) or independent of qualitative measures (Topping et al., 2003). However, none of these studies compare any gains in self-esteem in mathematics to the possible gains made by students not receiving tutoring services; these studies did not utilize control groups to provide comparisons. In addition, some studies indicated that future research should combine analysis of self-esteem and achievement in mathematics (Topping et al., 2003; Walker 2007).

Several studies exist regarding the effects of tutoring on both student achievement and self-esteem; the research that focuses on peer tutoring has potential for development. Barone and Taylor (1996) set up a peer tutoring model in which third grade students tutored first- and second-grade students by engaging in concrete activities. In preparation for interacting with tutees, the tutors were first introduced to a concrete activity, and then were asked to model the activity and teach it to a fellow tutor. The tutor partners then switch roles so the other tutor can practice. Tutors were then paired with a younger student throughout the course of thirteen days. Each day focused on a new activity. Both tutees and tutors wrote reflections in journals about their experience, and the feedback was positive from both sides. However, no quantitative instruments were used to collect data that would measure the effectiveness of such a model. Topping et al. (2003) reported significant gains in self-esteem during their five-week study as measured by the formal assessment administered to the tutee. However, the researchers reflect that a concurrent analysis of mathematics achievement would have been beneficial.

Duration of the study was a constraint for this study, as the researchers acknowledge that finding an appropriate instrument that could measure achievement gains after only five weeks of the tutoring intervention would have been difficult. In addition, the study did not set up a control group to determine whether the increases in self-esteem could be linked to the peer tutoring program.

There has been recent research that has conducted meta-analyses involving peer tutoring, but even those findings indicate that there are still areas to develop. In 2005, a meta-analysis conducted by Robinson, Schofield, and Steers-Wentzell, gathered previous studies in which students engaged in peer tutoring. The results were largely positive, citing that peer tutoring programs were effective in increasing self-esteem and academic achievement. The exception was instances in which the age gap was considerable; the researchers recommend caution when the difference in ages between tutor and tutee is that of six to seven years.

One area that was not developed in the meta-analysis was the effect of peer tutoring when implemented outside the school day as a supplement to the regular class instruction. The researchers note that tutoring was often conducted in place of teacher instruction, thereby reducing "the probability that the tutored subject gets additional time devoted to it" (Robinson, Schofield, & Steers-Wentzell, 2005, p. 353). In addition, no distinction was made in the meta-analysis between studies that did implement the tutoring program as a supplement to the teacher instruction and those that replaced teacher instruction, leaving the question of whether or not such an additional intervention would have any benefit (Robinson et al., 2005).

With so many varied studies focusing on peer tutoring, it is important to analyze their differences and similarities to this study. Current research indicates that there seem to be three key variables in implementation: length of tutoring (both sessions and programs), content of tutoring sessions, and structure of tutorial sessions (Baker et al., 2006; Barone & Taylor, 1996; Kroesbergen & Van Luit, 2003; Zuelke & Nelson, 2001; Walker, 2007).

Peer tutoring can vary in length and frequency, both in regards to individual sessions and overall programs. Some tutoring sessions may be as short as ten minutes, implemented as one of several stations during classroom activities, but occur daily (Dufrene, Noell & Gilbertson, 2005). In other studies, sessions may be held for forty minutes to an hour, yet occur fewer times a week, typically two to three times weekly (Fuchs et al., 2005; Zuelke & Nelson, 2001;). The longest duration of individual tutoring sessions appeared to be ninety minutes. In one study, these sessions were held several times a week (Walker, 2007), but in another study sessions of this length were limited to once a week (Baker et al., 2006).

The length of the tutoring program can vary as well. Many times program lengths will correlate to pre-existing time frames within the school calendar year, such as semesters or quarters (Walker, 2007), while others range generally from 16 to 20 weeks (Fuchs et al., 2008). Extremes on either end of the spectrum consist of programs that run for only five weeks (Topping et al., 2003) to a longer duration of eight months (Fuchs et al., 2005). Interestingly, a meta-analysis of tutoring studies by Kroesenbergen and Van Luit (2003) found that program length and effect size were negatively correlated, suggesting that tutoring efforts and content may be more focused in a more condensed amount of time. Ginsberg-Block, Rohrbeck, and Fantuzzo (2006) also found in their meta-analysis of peer-assisted learning programs that there was no significant difference between tutoring programs whose duration was less than 900 minutes in total and those programs whose duration was greater than 900 minutes in total. No reasoning was given as to why 900 minutes was significant.

Content, intent, and instructional methods of the tutoring program are other areas in which program differences are found. At times, the tutoring may act as a supplement to the mathematics curriculum used in the regular classroom (Fuchs et al., 2005). One study recommends that future peer tutoring interventions focus more on preparation and mastery of algebra (Fuchs et al., 2008). It also may take the form of addressing specific needs, such as homework help (Baker, Rieg, & Clendaniel, 2006; Walker, 2007), improving classroom performance (Zuelke & Nelson, 2001), testing preparation (Walker, 2007) or early intervention as a preventive measure (Fuchs et al., 2005). Kroesenbergen and Van Luit (2003) found that most tutoring studies were focused on basic skills and problem solving. Some tools used in tutoring programs involve structured lessons or scripts provided to peer tutors (Barone & Taylor, 1996; Fuchs et al., 2005), seat work such as problem sets (Dufrene, Noell, & Gilbertson, 2005), mathematical games intended to motivate students (Topping et al., 2003), or computer-assisted mathematics instruction (Mevarech & Rich, 1985). Walker (2007) suggests concrete, engaging activities. In a meta-analysis conducted by Ginsberg-Block, Rohrbeck, and Fantuzzo (2006), the researchers found that focusing the sessions to meet the needs of each tutee through individualized curriculum is another way to provide structure, though doing so did not necessarily increase self-concept more than tutorial programs with non-individualized curricula (Ginsberg-Block et al., 2006),

Structure is also recommended in the method of instruction utilized by tutors. Some research suggests that tutees benefit more from explicit, procedure-based instruction during tutoring sessions as opposed to more concept- or discovery-based instruction (Fuchs et al., 2008) but perhaps moving the instruction into more conceptual understanding as the program progresses and tutors become more confident (Walker, 2007).

While peer tutoring has been reported to have predominately positive effects on student achievement and attitudes towards mathematics, concerns have risen. One concern regarding such programs is the degree of organization and focus. In terms of organization, Dufrene, Noelle and Gilbertson (2005) implemented an organizational system that tracked progress using folders, worksheets for tutees, answer keys, and progress charts. However, they recommend the use of computerized programs to monitor gains and track corrective feedback.

Another concern is the accuracy of the tutors. Since the tutors are themselves students and are still learning, there is a higher incidence of errors in procedure and computation (Dufrene et al., 2005). Also, there is the possibility that tutors will not realize when the tutee has made an error, therefore missing a critical opportunity to redirect the tutee and help them correct themselves. One way of addressing this potential problem is to rank tutors and tutees by current achievement level or mathematics course and match them so the difference between tutor and tutee remains consistent (Topping et al., 2003). However, it is perhaps for this reason that in the meta-analysis of several tutoring studies, Kroesenbergen and Van Luit (2003) found that interventions that utilized peer tutoring were less effective.

Based on existing research, it is evident that though several studies exist that focus on peer tutoring, there are still several opportunities to investigate areas that have not received much attention. A study that focuses on the effects of peer tutoring on both academic achievement and attitudes about mathematics, as measured by quantitative instruments in comparison to a control group, specifically with middle school age tutees in the field of mathematics, would be a worthwhile addition to the current research.

Increased Learning Time

Another strategy, increasing instructional and support time by offering out-of-school time (OST) programs, is also widely practiced. Implementation of The No Child Left Behind Act has been a catalyst for the increase in development of these programs (Lauer, Akiba, Wilkerson, Apthorp, Snow & Martin-Glenn, 2006); schools that are not meeting Annual Yearly Progress (AYP) are required to provide evidence of supplemental instruction to students not meeting standards. While the visibility of OST programs has increased, there is a wide spectrum of implementation, purpose and structure.

The manner in which OST programs are implemented varies. One variable is the timeframe during which the program is executed. Some OST programs are held during the summer months, while others are held during the school year, but outside the regular school day. However, according to Lauer et al., the timeframe seems to have little impact on the programs' effectiveness; the impact of each was comparable and the differences were not statistically significant (304). Regardless, there were some differences in the content and focus for OST programs that were held during the school year versus during the summer programs. Summer programs often focused on non-academic areas in addition to academics. Many of these programs supplemented academic focus with more social and whole-child aspects, such as career planning and college readiness (Lauer et al., p. 298, 303).

These differences in program focus are not limited to summer OST programs in contrast with OST programs held during the school year; these differences are observed within similar timeframes as well. Regardless of the timeframe in which OST sessions are implemented, there is a wide spectrum of program goals. In a report that examines different types of OST programs, Miller (2003) cites that while many programs provide academic support to the attendees, in middle school OST programs in particular, the focus may be to reduce risky behaviors in teenagers, such as substance abuse, criminal activity, school attendance, conflict resolution, and problem solving (Miller, 2003, p. 53) In their meta-analysis of studies involving out-of school time (OST) interventions, Lauer et al. found that OST programs that do not solely focus on academic areas still can have positive effects on student achievement, but cautioned that administrators of OST programs should closely monitor the implementation of such programs in order to ensure that time is used as effectively as possible (2006).

Within the OST programs that focus on student academics, the daily routine may vary. Some OST programs utilize specific curricula, training their OST program staff to implement chosen curriculum. In a study conducted by Aidman (1997), teachers were provided with daily lesson plans that had been developed previously by administrators and summer school staff and corresponding necessary materials. The goals of the lesson plans were very focused, particularly in mathematics, which targeted six conceptual strands: number sense, operations, geometry, measurement, probability and statistics, and algebra. Each strand was taught for five consecutive days (Aidman, 1997).

As opposed to a concrete curriculum, OST programs may address the daily needs of the student, supplementing their regular school day by offering academic support. Some programs may resemble homework help sessions. A study conducted by Walker (2007) paired high school students with other high school students for homework help and test preparation during a drop-in tutorial class that was held after school. Similar programs have combined homework help with less structured games that reinforce students' mathematics skills (Baker et al., 2006).

The effects of OST programs on student achievement are generally positive, but there are several opportunities for further research. In their meta-analysis, Lauer et al. found that while both elementary and secondary students benefited from OST programs with content in reading and mathematics, secondary students tended to reap the benefits more consistently when the OST programs focused on mathematics (2006). However, due to the fewer numbers of secondary students participating in OST programs, the researchers note that further investigation is needed. In addition, the researchers recommend that specific details on program implementation and evidence of success be investigated and documented. Since OST programs are common interventions used by schools, identifying the defining characteristics of successful programs would influence future OST programs, increasing their own likelihood of success (Lauer et al., 2006). This proposed study was designed to contribute to this need.

Measuring Student Attitudes Towards Mathematics

Students' attitudes about mathematics have a strong impact on their achievement. Confidence in mathematics, including a sense of value in the subject and themselves, is essential to success. Most research involving students' attitudes about mathematics support the theory that those attitudes affect achievement, positively if the attitudes are themselves positive, negatively if the attitudes are negative (O'Neal et al., 1988; Steele & Arth, 1998). A study conducted by Topping et al. suggests that peer tutoring has a positive effect on attitudes about mathematics; however, the study did not have a control group against which to measure those attitudinal gains (2003). Without a control group, it is difficult to determine if the peer tutoring is responsible for the increases or if the increases would have occurred regardless of that intervention. The proposed study would measure the attitudes of both the experimental group and the control group, thereby making any connections between increased attitudes and peer tutoring more valid.

Several instruments exist for measuring mathematics attitudes, but the Fennema-Sherman Mathematics Attitudes Scales (FSMAS) (Fennema & Sherman, 1976) have proven over time to be effective. However, it is very long, taking on average 45 minutes for students to complete (Mulhern & Rae, 1998); therefore, some researchers have attempted to develop shortened, modified versions of the FSMAS that would not compromise its original validity. In one of the most successful studies by Mulhern and Rae (1998), the researchers administered the original 108-item FSMAS, then analyzed the results in order to best develop a shorter, 51-item assessment. All analyses suggest that the attempt was successful.

In light of all previously mentioned research, this study aims to answer the question: Will a combined intervention consisting of peer tutoring and out-of-school time contribute to greater increases in mathematics achievement and attitudes than those of students not exposed to the interventions?

CHAPTER 2: Method

Participants

The study was conducted at a charter school in St. Paul that serves grades 6-8 in the middle school and grades 9-12 in the high school. The intended tutees were seventh and eighth graders, while the intended tutors could range from seventh grade to twelfth grade. Students enrolled in the school's grade-level seventh grade mathematics courses (Math 7), standard Pre-Algebra courses, and a modified Pre-Algebra course (Math 8) were invited to join an after-school tutorial class in which they could opt to participate in the study as well. Students were invited to join the program through postings on teacher whiteboards and school announcements, then more formally on the after-school program registration sheets.

After the initial sign-up, any students who had not signed up but scored exceptionally low on the pretest were approached once more to give them one more chance to participate. Students could also join for the tutorial program regardless of their pretest scores. Students had a choice between signing up for tutoring two days a week or one day a week. After quarter grades were sent home and a potential need for additional mathematics support was more apparent, more students were admitted into the program. Some students were also placed in the program at the parents' request. There were a total of 34 students who participated as tutees in the program. Of these tutees, 22 attended one day a week and 12 attended both days.

Middle school tutors were recruited in a similar fashion: teacher postings, student announcements, and after-school program registration sheets. The high school math teacher and the counselor also approached high school tutors. Students at the high school are required to fulfill 20 volunteer hours per year in order to graduate, so the high school staff presented the program as a valuable way to earn hours. Middle school students and their families are asked to volunteer 50 hours per year, though there is no penalty for failing to fulfill all hours. Tutors could also sign up for two days a week or one day a week, though the two days option was encouraged for the sake of consistency. As a result, 22 students, ranging from seventh to twelfth grade, participated as tutors in the program. Of these 22 tutors, nine tutors attended sessions one day a week and the remaining 13 attended both days.

Materials and Procedure

Initial data collection. The study utilized pre-existing data from pretests administered at the start of the academic year. Pretests were administered by school staff in September for placement purposes and consisted of class material for the upcoming year as well as some review material. Each mathematics class utilized its own pretest that was relevant and appropriate to the content of the course and content from previous courses. The pretest was then scored and school staff made any necessary schedule changes.

In January, all students in Math 7, Math 8, and Pre-Algebra took a survey about their attitudes about mathematics. This survey was the shortened form of the FSMAS created in Mulhern & Rae's 1998 study, in addition to ten questions from a survey created and used in a study in a 1989 study by Alan H. Schoenfeld. Students were given forty-nine statements regarding several aspects of their mathematics experience in schools thus far. There was also one example question, which some students responded to and some did not; therefore the example question was not included in the scoring. While the survey was longer than desired, since the shortened FSMAS had been developed and tested in the Mulhern and Rae study (1988) and the last questions done so by Schoenfeld (1989), to tamper with the existing surveys would compromise the validity. The script for the directions, in addition to the first forty questions, including the example question, were adapted from the survey developed by Mulhern and Rae in 1998.

Like the original FSMAS, the items on the shortened FSMAS were coded and classified as one of seven different categories. There were five items about students' confidence in mathematics (C), eight items regarding the perceived usefulness (U) of mathematics in their lives, seven items assessing whether students felt as though mathematics is a male domain (MD), seven items dealing with students' perceived success (S) in mathematics, six items assessing the effectiveness and significance of the role of the mathematics teacher (T), five items regarding the students' level of anxiety (A) with mathematics, and one item about the students' effectance motivation (EM) and perseverance (Fennema & Sherman, 1976). Items were worded both positively and negatively. Because the Mulhern & Rae survey was originally administered to Irish students, the Mulhern and Rae survey was slightly edited for idiomatic language differences between the Irish-English and American-English vernacular. The remaining ten questions were adapted from a survey developed and administered for a 1989 study by Alan H. Schoenfeld. These questions address students' perceptions of the reasons they receive good and bad grades in mathematics courses. The full survey administered can be viewed in the Appendices.

For each of the fifty questions, the students were to indicate their level of agreement for each of the statements on a Likert-style scale on which 5 represented "strongly agree," 4 "somewhat agree," 2 "somewhat disagree" and 1 "strongly disagree." If the student was not able to answer the question or was not sure, they were instructed to choose 3. While it would have been preferable to design the scale with a forced choice, since both the shortened FSMAS and the Schoenfeld survey utilized this five-point scale.

To score the surveys, the positively-worded items were scored at face value and the negatively-worded items were inversely scored. For example, a negatively-worded statement that was answered with a 5 for "strongly agree" was scored as a 1 as though it had been a positively-worded statement. The scores for each item were then added for each student, resulting in the overall score. Then items were added within their coded categories, resulting in the categorical scores. Survey changes for overall scores and categorical scores alike were calculated by subtracting the pre-survey score from the post-survey score. A positive change indicates an increase in survey scores and thus a positive change in mathematics attitude; inversely, a negative change indicates a decrease in survey scores and thus a decline in mathematics attitudes.

Because data from all students, regardless of potential participation in the tutorial program, would be used for the purpose of creating baseline control group data and treatment group data, all students were told that their scores would be used confidentially and not shared, and were asked in writing to indicate whether or not they preferred their scores not to be used. If the students did not opt out, a similar form was sent home to parent(s)/guardian(s). No students or parents chose not to allow their data to be used.

Implementing the interventions. The peer tutoring sessions were held during a non-mandatory period (Enrichment) of the school day in which the majority of the students participate in various academic and non-academic activities. Participants had the option of attending one or two days a week for 40 minutes each session. Attendance was taken daily and hours were logged for both tutors and tutees. Assent forms were distributed to students who have elected to participate in the program. Consent forms were sent home to parents whose children are willing to participate in the study.

In order to match tutees and tutors, pretest scores were ranked within each group. Middle school tutors with high scores were matched with tutees that scored comparatively better than their fellow tutees. High school tutors were matched similarly, but their proficiency was based on teacher input and recommendations. Little consideration was given to gender pairings, though student age and grade was considered. All tutors were the same age or older than their tutees. Only one pair consisted of students who were in the same mathematics course. Some tutors that were exceptionally proficient in mathematics as indicated by pretest scores or teacher input worked with two tutees that were in the same mathematics class.

Once tutees were registered for the program, they were required to attend a one-hour training session staffed by the teachers who would supervise the tutorial program. Training consisted of identifying characteristics of good tutors, becoming familiar with the needs of the tutee, guidelines for successful tutoring sessions, and suggested procedures for helping students with mathematics. Tutors were then shown a tutorial log that would be used to keep record of the sessions, including the date of the session, the tasks that were attempted (homework, studying, general help, etc.), the mathematical content of the tasks, and a reflection on the perceived success of the task. Tutors were given a folder for each of their tutees that would hold the tutorial logs for future use. All folders were kept with the supervisors and brought to all sessions.

Tutoring sessions were held in two places throughout the study. During the first half of the program, the sessions were held in a mathematics classroom. A seating chart was given, placing tutors and tutees next to one another and as far away from potential distractions such as friends and other students as possible. However, this classroom only had 30 desks, so during the peak sessions in which all scheduled tutors and tutees were present, there would not be sufficient seating or room for students to work. Therefore, for the second half of the program, the sessions moved to the school cafeteria, which gave tutoring pairs ample room to spread out and work without being as distracted.

Once IRB approval was attained, the program sessions began in January and proceeded through June. During this time, tutors helped tutees with homework, studying for tests and quizzes, and problems from the Accelerated Math program by Renaissance Learning, which is an instructional management system. The Accelerated Math program allows instructors to assign problems from chosen objectives to individual students, thereby creating an individual curriculum for each student. Students can work at their own pace towards mastery. The supervisors often had access to computers and could score students' Accelerated Math work immediately. If a student scored well enough (80% proficient), they could master the objective by taking a test without the help of the tutor. If they passed, they were assigned new objectives. If not, they were given more practice on the current objectives and the tutors would provide additional assistance.

Final data collection. At the end of the school year, students were administered both the academic posttest for their mathematics course and the attitudinal survey again. The posttests were made by evaluating each pretest from the beginning of the year, deleting items about topics not addressed over the course of the year. So that each course's posttest was identical to its pretest, the same items were deleted from the pretests and pretest scores were adjusted to reflect any changes made to the posttest to account for deleted items. After the pretest scores were modified in order to match the items that had been on the posttest and not the items that had been removed, the pretests and posttests were otherwise scored identically. Scores for both tests were scaled to 100 total points. The attitudinal survey given at the end of the year was identical to the initial survey.

As with the academic data, the results of the pre-survey and post-survey were added to the data file in SPSS containing the pretest and posttest data. Any missing values were assigned “999,” which was the designated placeholder and would not affect the statistics. From that data, the point change in each student’s survey was calculated. In addition, the survey data was also categorized by the codes mentioned previously: Confidence (C), Usefulness (U), Male Domain (MD), Success (S), Teacher (T), Anxiety (A), Effectance Motivation (EM), and the last ten questions from Schoenfeld’s 1989 survey (Schoen). Presurvey scores, postsurvey scores, and changes between the two scores for each category were calculated using SPSS.

Preparation for data analysis. After scores for the academic tests and surveys were calculated, each student was assigned a three-digit code, thereby removing their names from the all data. Identification numbers, grade level, mathematics class, whether they had received tutoring services were then entered into SPSS for tracking and analysis. For the students who had received tutoring services, the number of service hours received was recorded as well; for students who did not receive services, this number remained 0. All pretest, posttest, pre-survey, and post-survey scores were entered as well.

Data for 99 students were used for the academic portion of the study, which differs from the *n* for the attitudinal survey due to mathematics placement changes from the initial placement to the final placement. When the students are scheduled for their mathematics classes at the beginning of the year, those placements are subject to change based on observed success in the current math class and any supplementary information that can be gleaned from student records (which tend to arrive at new schools later than placement is made) such as the individualized education programs (IEPs) previous math courses, corresponding grades, and past performances on standardized tests. Because the pretest and posttests, while the same measure of knowledge, differ between math courses (Math 7, Math 8, and Pre-Algebra), if a student took a pretest in one course, was subsequently moved to a different course, and took the posttest for that course because it better matched the content covered, they were not included in the pre-posttest analysis.

As noted before, the number of students who participated in the attitudinal survey is greater than that of the achievement tests. Some scores are missing due to absenteeism during one or more days the survey was given. There are a total of 112 students in the study; 99 students participated in the achievement portion of the study and 103 participated in the attitudinal survey. Ninety students participated in both portions of the study.

CHAPTER 3: Results

In order to determine if the research question of whether peer tutoring and out-of-school time in mathematics would positively affect academic achievement and attitudes towards mathematics, data was analyzed using SPSS, and the findings were interpreted.

Pretests and Posttests

All pretest and posttest data were first disaggregated according to whether they received tutorial services or not (“Yes” versus “No”). The control group experienced a mean means, standard deviations, and medians are given in Table 1. When comparing the results between the two groups, the means of both the pretest and posttest are higher for the control group than the treatment group. The growth in percentage points between pretest and posttest for both groups is comparable. The mean in difference in percentage points is slightly lower (34.07 versus 36.04, *SD* = 12.825, 11.30, respectively) for the

Table 1
Academic Pre- and Posttest Results

Services		Pretest	Posttest	Δ% Pt	% Increase
No	Mean	34.48	70.52	36.04	124.19
(*n* = 69)	*SD*	12.441	14.500	11.299	73.869
	Median	33.00	72.00	36.00	111.76
Yes	Mean	20.70	54.77	34.07	198.41
(*n* = 30)	*SD*	9.473	16.322	12.825	115.314
	Median	19.50	58.00	36.50	159.93
Total	Mean	30.30	65.75	35.44	146.68
(*N* = 99)	*SD*	13.208	16.664	11.751	94.321
	Median	28.00	68.00	36.00	124.24

students who received services, whereas the median is slightly higher (36.5 versus 36).

When posttest results are compared exclusively to the corresponding students' pretest results and percent increase is calculated, the results differ. The control group experienced a mean percent increase of 124.19 (*SD* = 73.87) and a median percent increased of 111.76, while the treatment group displayed a mean percent increase of198.41 (*SD* = 115.13) and a median percent increase of 159.93. Given, the pretest scores for the treatment group were typically lower than those of the control group—some scores as low as 7%. Thus, the percent increases for the treatment would tend to be higher. Therefore, this statistic was not analyzed further.

In order to determine whether the mean percentage point increases between the two groups were significantly different, two analyses were conducted. First, an independent sample *t*-test (*CI* = 95%) was conducted to examine the difference between the means of the percentage point differences between the treatment group and the control group. Given the results of the test, there does not appear to be a significant difference between the two means, $t(97) = -0.77$, $p = .45$. Second, the two means were compared by calculating the effect size. Cohen's *d* and Pearson's *r* were calculated using a web-based effect size calculator (http://www.uccs.edu/~faculty/lbecker/), which utilizes means and standard deviations. Effect sizes were then interpreted using a scale hosted by the same website (http://www.uccs.edu/~faculty/lbecker/es.htm). The treatment group represented the first group and the control group represented the second group. The resulting *d* value was -0.16 and the *r* value was -.08, indicating a low effect size.

A series of Pearson's correlation analyses were calculated between students' pretest and posttest scores to determine whether students with higher pretest scores would have higher posttest scores. The bivariate correlation between pretest and posttest is positive, $r(97) = .71$, $p < .001$. In order to determine whether the presence of tutoring services had any effect on the relationship between pretests and posttests, I did a partial correlation analysis. When controlled for whether or not students received services, the *r* value decreases to 0.64 with $p < .05$. When instead controlling for number of hours tutored, the *r* value of 0.67 is still lower than the bivariate correlation but is higher than the partial correlation *r* controlling simply for whether or not students received services.

Correlation coefficients were also calculated to determine the potential correlation between presence of services and percentage point differences between pretests and posttests. While it appears to be a slightly negative correlation ($r = -.08$) the *p* value of 0.45 indicates that this correlation is not significant. All results of the correlation analysis can be found in Tables 14, 15, and 16 in the Appendix.

Table 2
Academic Pre- and Posttest Results Analyzed by Mathematics Class

Math Class	Services	Stat.	Pretest	Posttest	% Pt Difference
Math 7	No (*n* = 31)	*M*	43.00	75.55	32.55
		SD	11.48	14.24	9.78
		Median	41.00	77.00	31.00
	Yes (*n* = 11)	*M*	25.91	60.18	34.27
		SD	9.995	12.327	9.655
		Median	25.00	59.00	35.00
Math 8	No (*n* = 2)	*M*	21.00	45.00	24.00
		SD	.000	2.828	2.828
		Median	21.00	45.00	24.00
	Yes (*n* = 8)	*M*	12.00	41.00	29.00
		SD	5.831	19.530	15.446
		Median	10.50	42.50	34.00
Pre-Algebra	No (*n* = 36)	*M*	27.89	67.61	39.72
		SD	8.28	12.89	11.52
		Median	28.5	70.00	38.50
	Yes (*n* = 11)	*M*	21.82	59.36	37.55
		SD	6.69	12.10	13.52
		Median	24.00	58.00	40.00
Total	No (*n* = 69)	*M*	34.48	70.52	36.04
		SD	12.44	14.50	11.30
		Median	33.00	72.00	36.00
	Yes (*n* = 30)	*M*	20.70	54.77	34.07
		SD	9.47	16.32	12.83
		Median	19.50	58.00	36.50

The statistics were then calculated with respect to services and mathematics class. These results are displayed in Table 2. As with the collective results, the control group had higher pretest and posttest means and medians. However, attention was focused on the percentage point increases. The treatment groups in Math 7 and Math 8 experienced higher mean percentage point increases than the students in the control group. Students in Math 7 in the control group (n = 31) experienced a mean percentage point increase of 32.55 (SD = 9.78) compared to the treatment group in Math 7 (n = 11) who experienced an increase of 34.27 (SD = 9.66). Students in Math 8 demonstrate similar results, with the control group (n = 2) showing a mean percentage point increase of 24.00, (SD = 2.83) and the treatment group (n = 8) showing a mean increase of 29.00 (SD = 15.45). Pre-Algebra, however, had the opposite results. The control group's mean (n = 36) was 39.72 (SD = 11.52), higher than the treatment group's mean (n = 11) of 37.55 (SD = 13.52). When comparing each class' medians, the treatment groups' medians were higher than the control groups' in all three classes: Math 7 (35.00 compared to 31.00), Math 8 (34.00 compared to 24.00), and Pre-Algebra (40.00 compared to 38.50).

To help determine whether the differences between the mean percentage point increases in each class were significant, two main analyses were executed. First, the data was sorted into mathematics classes and each class' data was copied into three separate files. An independent samples t-test was then performed on each of the three databases in order to determine if the mean percentage point differences were significant for each class. In Math 7, the means were not considered to be significantly different, $t(40) = 0.50$, $p = .62$. The results were similar for Math 8, $t(8) = 0.44$, $p = .67$. For the Pre-Algebra class, the means were not significantly different either, $t(45) = -0.53$, $p = .60$. Second, effect sizes were calculated. For Math 7, Cohen's d was calculated to be 0.18

and *r* was calculated to be 0.09, indicating a small effect size. For Math 8, *d* was calculated to be 0.45 and *r* was 0.22, which represents a medium effect size. For Pre-Algebra, Cohen's *d* was -0.17 and *r* was -0.09, indicating a small but negative effect size between the treatment group mean and the control group mean.

The final analysis for the academic data was to determine whether there was a correlation between the number of hours a student was tutored and the academic results. First, the data was split into the control group and the treatment group. Then, Pearson correlation coefficients were calculated for number of hours tutored compared to the percentage point difference for each student in the treatment group. The results in Table 3 suggest that there was a slight correlation, that correlation was not very significant, $r(28) = .248$, $p = .187$.

The data was then split by mathematics classes and the correlation analysis was repeated for the treatment groups within each class. While Math 7 and Math 8 show no significant correlations between number of hours tutored and posttest scores or percentage point differences, the Pre-Algebra class does. The positive correlation

Table 3
Correlations between Hours Tutored and Academic Percentage Point Differences

		Hrs	Percentage Point Difference
Hrs	Pearson Correlation	1	.248
	Sig. (2-tailed)		.187
	N	34	30
Percentage Point Difference	Pearson Correlation	.248	1
	Sig. (2-tailed)	.187	
	N	30	30

Table 4
Correlations between Hours Tutored, Academic Posttests, and Percentage Point Differences

Class		Hrs	Posttest	Percentage Point Difference
Math 7 Hrs	Pearson Correlation	1	-.260	-.422
	Sig. (2-tailed)		.439	.196
	N	13	11	11
Math 8 Hrs	Pearson Correlation	1	.330	.186
	Sig. (2-tailed)		.424	.659
	N	10	8	8
Pre-Algebra Hrs	Pearson Correlation	1	.708*	.717*
	Sig. (2-tailed)		.015	.013
	N	11	11	11

*. Correlation is significant at the 0.05 level (2-tailed).

between hours tutored and posttest scores appears to be significant ($r(9) = .708$, $p = .015$), as does the positive correlation between hours tutored and percentage point difference between the pretest and posttest, $r(9) = .717$, $p = .013$. Complete results are displayed in Table 4.

Presurvey and Postsurvey

Using SPSS, analysis of the survey results began first by calculating descriptive statistics. With a high score of five for each of the 49 questions, the maximum score possible was 245. The mean presurvey score ($N = 103$) was 195.56 ($SD = 21.83$), median of 199.00, maximum score of 244 and a minimum score of 100. The mean of the postsurvey scores ($N = 103$) was 197.58 ($SD = 23.84$), median of 202.00, maximum of 236 and minimum of 105. The mean point change between the presurvey and postsurvey was 2.02 ($SD = 17.14$), median of 1.00, maximum of 36 and minimum of -87. The minimum point change was a considerable outlier, indicating a large decrease in attitude towards mathematics. A number of factors could contribute to this outlier such as confusion with ranking scale, but no follow up was done to determine the reasons.

The presurvey and postsurvey means are so close together that I performed a paired samples *t*-test to determine whether or not these two means were significantly different. First, there appears to be a strong positive correlation between a student's presurvey and postsurvey scores ($r = .722$, $p < .001$). However, the paired-sample *t*-test suggests that there is not a significant difference between an individual's presurvey and postsurvey, $t(102) = -1.196$, $p = .235$.

When the data is analyzed in light of whether students received tutoring services, however, slight differences are revealed. The survey data was separated into the control group and treatment group, at which point the mean presurvey, postsurvey, and survey point changes were calculated and compared between the two groups. The results are displayed in Table 5. Students who did not receive tutoring services (n = 73) show a presurvey mean of 198.96 points (*SD* = 22.37) and median of 202.00, postsurvey mean of 200.18 (*SD* = 21.28) and median of 203.00, and survey change mean of 1.22 points (*SD* = 14.56) and median of -1.00. Students who did receive tutoring services (n = 30) show a

Table 5
Attitudinal Survey Results with Respect to Services

	Services					
	No (*n* = 73)		Yes (*n* = 30)		Total (*N* = 103)	
	M	*SD*	*M*	*SD*	*M*	*SD*
Survey-Pre	198.96	22.37	187.30	18.29	195.56	21.83
Survey-Post	200.18	21.28	191.27	28.56	197.58	23.84
Survey Change	1.22	14.56	3.97	22.40	2.02	17.14

presurvey mean of 187.30 (*SD* = 18.29) and median of 189.00, postsurvey mean of 191.27 (*SD* = 28.56) and median of 193.00, and survey change mean of 3.97 (*SD* =22.40) and a median of 9.50.

As with the academic data, two methods were used to determine whether the two survey change means were significantly different. First, I executed an independent-samples *t*-test. Based on the results, it does not appear that the survey change means between the two groups differ by a significant amount, $t(101) = 0.74$, $p = .46$. Then, I calculated effect size. Cohen's *d* was 0.14 with an *r* of 0.07, both of which suggest a small effect size.

After means for control and treatment groups were calculated and analyzed, data was disaggregated with respect to mathematics class. When the survey changes are compared, Math 7 ($n = 41$) demonstrated a mean increase of 2.73 points (*SD* = 21.21) and a median increase of 3.00 points. Math 8 ($n = 13$) demonstrated a mean increase of 3.08 points (*SD* = 15.90) and a median increase of 3.00 points. Pre-Algebra students ($n = 49$) demonstrated a mean point increase of 1.14 (*SD* = 13.59) and a median point change of 0.00.

When the survey change results are compared within classes with respect to whether or not students received services, the mean increase for students in Math 7 in the control group ($n = 29$) was 3.00 points (*SD* = 15.79) and the median increase was 1.00. Math 7 students in the treatment group ($n = 12$) demonstrated a slightly lower mean of 2.08 (*SD* = 31.64), but a significantly higher median change of 11.00 points. Math 8 showed increases for the treatment group (*n* = 9) for both the mean point increase and the median point increase with 3.44 and 8.00, respectively (*SD* = 13.94). This is higher than the mean and median point increases for the control group (*n* = 4), which were 2.25 and 0.50, respectively (*SD* = 22.17). Pre-Algebra followed the same trend with the control group (*n* = 40) showing a mean decrease of -0.17 points in survey change (*SD* = 13.04) and a median decrease of -1.50. Contrast these results with the Pre-Algebra students in the treatment group (*n* = 9) who demonstrate a mean increase of 7.00 points (*SD* = 15.25) and a median increase of 9.00 points. Complete results are shown in Table 6.

The mean survey changes of the control groups and treatment groups within each class were then analyzed. Independent-samples *t*-tests were conducted for each mathematics class. For Math 7, the difference between the control group mean and the

Table 6
Attitudinal Survey Change with Respect to Services and Mathematics Class

Services	Mathematics Class	*M*	N	*SD*	Median
No	Math 7	3.00	29	15.79	1.00
	Math 8	2.25	4	22.17	.50
	Pre-Algebra	-.17	40	13.04	-1.50
	Total	1.22	73	14.56	-1.00
Yes	Math 7	2.08	12	31.64	11.00
	Math 8	3.44	9	13.94	8.00
	Pre-Algebra	7.00	9	15.25	9.00
	Total	3.97	30	22.40	9.50
Total	Math 7	2.73	41	21.21	3.00
	Math 8	3.08	13	15.90	3.00
	Pre-Algebra	1.14	49	13.59	.00
	Total	2.02	103	17.14	1.00

Table 7
Attitudinal Presurvey Scores by Category

	Services					
	No (n = 73)		Yes (n = 30)		Total (N = 103)	
	M	*SD*	*M*	*SD*	*M*	*SD*
Confidence	18.16	4.61	15.67	4.80	17.44	4.78
Usefulness	33.64	5.85	31.50	6.66	33.02	6.15
Male Domain	33.36	2.50	32.37	3.06	33.07	2.69
Success	28.88	4.30	28.47	3.45	28.76	4.06
Teacher	23.62	4.67	23.50	3.50	23.58	4.34
Anxiety	18.14	4.61	15.93	4.32	17.50	4.61
Effectance Motivation	3.49	1.02	3.33	1.32	3.45	1.11
Schoenfeld	39.67	5.28	36.53	4.87	38.76	5.34

treatment group mean was not at all significant, $t(39) = -0.12$, $p = .90$. Nor were the differences between the means significant for Math 8, $t(11) = 0.12$, $p = .91$. The differences in means for the Pre-Algebra treatment and control groups were moresignificantly different than those in the other two classes, but still not enough to be conclusive, $t(47) = 1.45$, $p = .16$. Then, effect sizes were analyzed for each class. Cohen's *d* was small for both Math 7 and Math 8, -0.04 and 0.06, respectively. The effect size *r* was small for each as well; Math 7 had an *r* of -0.02 while Math 8's *r* was 0.03. This was not the case for Pre-Algebra, however, whose *d* was calculated to be 0.51 and whose *r* was 0.24, indicating a medium effect size.

All the data was then split into the eight different categories, seven of which are categories defined by the Fennema-Sherman Mathematics Attitudes Scale: Confidence, Usefulness, Male Domain, Success, Teacher, Anxiety, and Effectance Motivation. The eighth category was adapted from the 1989 Schoenfeld study and addresses the perceived circumstances that contribute to a student's success in mathematics. Each survey question had been coded as one of the eight categories and the presurvey, postsurvey, and survey change totals were compiled for each category. The highest points possible for each survey are as follows: Confidence-25, Usefulness-40, Male Domain-35, Success-35, Teacher-30, Anxiety-25, Effectance Motivation-5, Schoenfeld-50. All presurvey scores were higher for the control group than the treatment group, as shown in Table 7. However, while several of the mean postsurvey scores were still higher for the control group, the treatment group mean in the Success and Teacher categories surpassed those of the control group. The complete results of the postsurvey analysis are shown in Table 8.

Perhaps more telling is the mean point changes between the surveys for each group. Though the control group maintained higher means on six of the eight categories,

Table 8
Attitudinal Postsurvey Scores by Category

	Services					
	No (*n* = 73)		Yes (*n* = 30)		Total (*N* = 103)	
	M	*SD*	*M*	*SD*	*M*	*SD*
Confidence	18.38	4.39	16.90	4.18	17.95	4.36
Usefulness	33.37	6.30	32.00	7.51	32.97	6.67
Male Domain	33.67	2.38	31.37	5.87	33.00	3.86
Success	28.71	4.01	29.30	4.07	28.88	4.02
Teacher	24.00	4.31	24.07	4.97	24.02	4.49
Anxiety	18.25	4.85	17.53	4.50	18.04	4.74
Effectance Motivation	3.71	.98	3.43	1.28	3.63	1.08
Schoenfeld	40.08	4.68	36.67	6.51	39.09	5.48

the treatment group had higher mean point changes in all but two of the categories. The only two categories that the treatment group mean point changes did not exceed those of the control group were Male Domain and the Schoenfeld categories. All results are shown in Table 9.

After seeing the mean survey changes for each category, I compared the two groups' means using independent-samples t-tests and effect sizes. First, I conducted a series of independent-samples *t*-tests to determine whether the differences were significant. All results are shown in Table 17, located in the Appendix. Based on the data produced, it does not appear that any particular category had a significant difference between the means. I then calculated effect sizes for each category. Based on Cohen's *d* and effect sizes *r* for each category, the only category with a medium effect size is the Anxiety category, which had a *d* value of 0.38 and an *r* value of .19. Complete results of the effect size analysis are recorded in Table 10.

Table 9
Attitudinal Survey Score Changes by Category

	Services					
	No (*n* = 73)		Yes (*n* = 30)		Total (*N* = 103)	
	Mean	Std. Deviation	Mean	Std. Deviation	Mean	Std. Deviation
Confidence	.22	6.35	1.23	3.79	.51	5.73
Usefulness	-.27	4.13	.50	7.67	-.05	5.37
Male Domain	.32	2.22	-1.00	5.59	-.07	3.57
Success	-.16	3.65	.83	4.88	.13	4.05
Teacher	.38	3.78	.57	4.07	.44	3.85
Anxiety	.11	3.93	1.60	3.90	.54	3.96
Effectance Motivation	-.22	1.39	-.10	1.45	-.18	1.40
Schoenfeld	.41	3.80	.13	5.78	.33	4.44

Table 10
Effect Sizes for Means for Each Attitudinal Survey Category

Category	**Cohen's *d***	**Effect Size *r***
Confidence-Change	.19	.10
Usefulness-Change	.13	.06
Male Domain-Change	-.31	-.15
Success-Change	.23	.11
Teacher-Change	.05	.02
Anxiety-Change	.38	.19
Effectance Motivation-Change	.08	.04
Schoenfeld-Change	-.06	-.03

As with the survey data previously, the categorical data was separated into the three different math classes. Mean survey point changes were then calculated and compared between the treatment groups and control groups. For Math 7, the treatment group had higher mean survey changes in Usefulness, Success, and Anxiety. For the remaining categories, the control group means were higher. For Math 8, the treatment group's means were higher in seven of the eight categories, all except Anxiety. For the Pre-Algebra class, the treatment group's mean changes are higher in all eight categories. Complete results are reported in Table 11.

To determine if these differences were significant, independent-samples *t*-tests were conducted and effect sizes were calculated. In Math 7, none of the mean changes in the eight categories were significant, though the mean difference for the Anxiety category was somewhat close to significant, $t(39) = 1.74$, $p = .091$. In addition, most effect sizes for Math 7 are small; the only categories suggesting a medium to large effect size are

Table 11
Mean Attitudinal Survey Changes by Category with Respect to Services

Class	Services		C	U	MD	S	T	A	EM	Schoen.
Math 7	No	*M*	-.17	-.93	.41	.31	.72	.03	-.48	1.41
	(*n* = 29)	*SD*	6.11	4.28	1.62	3.74	4.85	4.44	1.41	3.09
	Yes	*M*	-.83	-.33	-2.50	.50	.50	2.50	-.67	-.25
	(*n* = 12)	*SD*	3.41	10.11	8.42	6.61	4.70	3.23	1.67	7.12
	Total (*n* = 41)	*M*	-.37	-.76	-.44	.37	.66	.76	-.54	.93
		SD	5.42	6.40	4.81	4.67	4.75	4.24	1.47	4.61
Math 8	No	*M*	.50	.25	.25	-1.00	-1.25	2.75	-1.00	-.50
	(*n* = 4)	*SD*	4.73	4.11	1.50	5.16	1.89	5.38	2.16	6.61
	Yes (*n* = 9)	*M*	3.67	.67	.56	.67	-.22	.33	-.11	-.22
		SD	1.80	5.48	1.59	3.94	3.46	5.03	1.36	5.14
	Total (*n* = 13)	*M*	2.69	.54	.46	.15	-.54	1.08	-.38	-.31
		SD	3.17	4.93	1.51	4.20	3.02	5.04	1.61	5.35
Pre-Algebra	No	*M*	.48	.15	.25	-.42	.30	-.10	.05	-.23
	(*n* = 40)	*SD*	6.77	4.06	2.65	3.49	2.98	3.37	1.26	3.89
	Yes	*M*	1.56	1.44	-.56	1.44	1.44	1.67	.67	1.00
	(*n* = 9)	*SD*	4.45	6.37	2.65	3.17	4.04	3.54	.87	4.85
	Total	*M*	.67	.39	.10	-.08	.51	.22	.16	.00
	(*n* = 49)	*SD*	6.38	4.52	2.64	3.48	3.18	3.43	1.21	4.06

Note. C = Confidence, U = Usefulness, MD = Male Domain, S = Success, T = Teacher, A = Anxiety, EM = Effectance Motivation, Schoen. = Schoenfeld

Male Domain, Anxiety, and the Schoenfeld questions. The Male Domain and Schoenfeld categories actually show a negative medium effect size. For Male Domain, Cohen's *d* was -0.48 and effect size *r* was -0.23. For the Schoenfeld questions, Cohen's *d* was -0.30 and *r* was -0.15. However, the Anxiety category showed a medium to large effect size in the positive direction; *d* was 0.64 and *r* was 0.30. Complete results of the *t*-tests and effect size calculations are displayed in Tables 18 and 21 in the Appendix. In Math 8, all results of the *t*-tests suggest the mean differences are not significant. However, the effect sizes indicate otherwise, particularly in the Confidence category, (*d* = .89, *r* = .40). In addition, medium effect sizes were calculated for the Success category (*d* = .36, *r* = .18), the Teacher category (*d* = .37, *r* = .18) and the Effectance Motivation category (*d* = .49, *r* = .24). However, the effect size for the Anxiety category was considered medium but negative (*d* = -.46, *r* = -.23). Complete results are reported in Tables 19 and 21 in the Appendix.

The Pre-Algebra classes showed similar results. The results of the independent-samples *t*-tests indicated that none of the mean differences within any category were significant. However, the effect sizes in some categories were found to be somewhat significant. In the categories of Success (*d* = .52, *r* = .25), Teacher (*d* = .32, *r* = .16) and Anxiety (*d* = .51, *r* = .28), the effect sizes were considered medium. On the Schoenfeld items (*d* = .28, *r* = .14), the effect size was classified as small to medium. However, the Male Domain category displayed a medium but negative effect size (*d* = -.31, *r* = -.15). Comprehensive results are recorded in Tables 20 and 21 in the Appendix.

As with the academic data, the final survey data analysis was to determine if there was a correlation between the number of hours a student was tutored and the point changes between their overall presurvey and postsurvey scores. Only the treatment group was considered in this analysis. The correlation was only slightly positive, but not significant, *r*(28) = .060, *p* = .752. Very few of the survey changes within the eight categories displayed a significant correlation, either positive or negative. The two exceptions were the Anxiety category and the Schoenfeld items. The correlation between hours tutored and score changes in the Anxiety category was nearly significant, *r*(28) = .348, *p* = .060. Similarly, the Schoenfeld items also have a positive correlation that is close to being significant, *r*(28) = .325, *p* = .079. Complete results are displayed in Table 12.

The same correlation analysis was conducted with the data disaggregated by mathematics class as well. The findings were similar to those of the overall treatment group, with a few exceptions. While the none of the correlations in the Math 7 or Math 8 classes are significant, the Pre-Algebra class showed some nearly significant correlations in the overall survey point change (*r*(7) = .627, *p* = .071) as well as in the Anxiety category (*r*(7) = .594, *p* = .092). Table 13 displays the complete results of the analysis.

Table 12
Correlations between Hours Tutored and Attitudinal Survey Changes

	Hrs		
	Pearson Correlation	Sig. (2-tailed)	N
Hrs	1		34
Survey Change	.060	.752	30
Confidence-Change	.099	.601	30
Usefulness-Change	-.083	.664	30
Male Domain-Change	-.118	.533	30
Success-Change	-.074	.696	30
Teacher-Change	.006	.977	30
Anxiety-Change	.348	.060	30
Effectance Motivation-Change	-.075	.695	30
Schoenfeld-Change	.325	.079	30

Table 13
Correlations between Hours Tutored and Attitudinal Survey Changes by Class

Math Class		Hrs	Overall	C	U	MD	S	T	A	EM	Schoen
Math 7 Hrs	*r*	1	-.28	.17	-.40	-.50	-.45	-.18	.42	-.17	.26
	Sig. (2-tailed)		.38	.59	.198	.10	.14	.59	.17	.60	.42
	N	13	12	12	12	12	12	12	12	12	12
Math 8 Hrs	*r*	1	.22	.07	.02	.32	.45	-.10	.20	-.23	.35
	Sig. (2-tailed)		.57	.86	.97	.41	.22	.81	.60	.55	.36
	N	10	9	9	9	9	9	9	9	9	9
Pre-Algebra Hrs	*r*	1	.63	.16	.26	.52	-.25	.28	.59	.11	.47
	Sig. (2-tailed)		.07	.69	.50	.15	.52	.47	.09	.78	.20
	N	11	9	9	9	9	9	9	9	9	9

CHAPTER 4: Discussion

Based on the results and analyses, it is difficult to state definitively whether the interventions of peer tutoring and increased time with mathematics positively affected student achievement or attitudes about mathematics. The data tells a few different stories, depending on the group and the analysis. When considering the academic data, the differences between pretest scores and posttest scores were very similar for the control group and the treatment group; in fact, when comparing the mean percentage point differences between the two groups for all students, each analysis suggests that any discrepancy is not significant. Since the treatment group had lower pretest scores than the control group, it follows from this that the gap between the two groups did not significantly close; the treatment group scores for the posttest are still lower than those of the control group by roughly the same amount. However, one can also conclude that while the achievement gap between the two groups did not close, nor did it widen. Without the interventions, the treatment group could have fallen further behind the control group. In order to determine whether this would have happened, a future study could be conducted in which the treatment and control groups are randomly assigned.

When the academic results were disaggregated with respect to mathematics class, the story changes a bit. For Math 7 and Math 8, the centers of data were both higher for the treatment group. With Pre-Algebra, the mean for the treatment group was lower but the median was higher. Pre-Algebra contributed the most student data for the control group, but had the smallest percentage of students participating in the study. Had the number of Pre-Algebra students in the treatment group been more commensurate with the participants in the other two classes, the results may have been different. Regardless, none of the tests for significance suggested a strong difference between the control groups and treatment groups, perhaps with the exception of the effect size result from the Math 8 class. With a medium effect size, one might interpret that slightly higher percentage point difference as partly a result of the interventions. However, because there does not appear to be a clear *yes* or *no* for any one group, the results are inconclusive.

When the survey data was examined, one of the first observations was that the presurvey and postsurveys were not very different from one another regardless of the group considered. In fact, when taken as a whole group regardless of services received, the pre- and postsurvey means were so similar that a paired-sample *t*-test did not even detect a significant difference. This may indicate that the students are exhibiting signs of consistency as it pertains to their opinions about mathematics.

There is a slight difference between the centers of data when the groups are split between the control and treatment. The mean survey changes for the two groups are only within three points of each other (out of a total possible score of 245 points), so it is not surprising that neither *t*-test nor effect size deem the difference significant. However, there is a more noticeable difference between the median point changes; the treatment group median was over ten points higher than the control group median. One contributing factor to the discrepancy between the treatment group's mean and median could be the outlier seen in the survey results. This discrepancy between the differences in mean survey changes and the median survey changes carried through to the class analysis as well, with the exception of the Pre-Algebra class. Both survey change means and medians were significantly higher than those of the control group, and the nearly significant *t*-test and medium effect size suggest it might not be a coincidence.

For the survey categories, the data suggests that, while there may have been some differences between the treatment groups and the control groups, the *p* values were so high in the majority of the cases that those differences are rendered moot. The effect sizes were small to medium for most of the categories, and in the class where the largest effect size was observed, the sample size was the smallest. There were a couple of places where some confidence in the results could be seen, namely with the effect sizes for the Anxiety category in Math 7, the Confidence category in Math 8, and the Success category in Pre-Algebra, indicating an increase in attitudes towards mathematics in theses areas. With more time, these results could be examined further.

Perhaps one of the more intriguing areas of analysis was with the correlations between the results and the number of hours the treatment group participants were tutored. The Pre-Algebra group's pretest and posttest data suggested a strongly positive, significant correlation—that the more hours students were tutored, the higher their posttests and percentage point differences were. None of the other correlations were that clear in either direction, indicating that in the other classes and with the survey results, it did not matter how often a student was tutored; perhaps just participating was as effective.

This study was not without its limitations. First, while one semester sounds like plenty of time to implement interventions, because some students only attended one day a week for forty minutes, the program may not have been used to its fullest potential. Frequent absenteeism for some students made it difficult for some tutors and tutees to build relationships and trust. Many last minute changes in tutoring pairs occurred, which did not help establish a routine. Space limitations were also an issue; the first tutorial room was extremely overcrowded, oftentimes students did not have chairs or could not hear one another against the surrounding tutor pairs. The second tutorial room was the school cafeteria, and while there was ample room, there was a noticeable energy shift. In the first room, the students were in their regular mathematics classroom, in an environment that promotes and enforces academic focus. The cafeteria, on the other hand, is exactly the opposite. Students were used to socializing in the cafeteria, and it was thus more difficult to get students to stay on task. Also, because the cafeteria was in the basement of a thick-walled building, the connection to the Internet was sporadic at best. Therefore, if tutees finished their homework and their Accelerated Math before the session was complete, it was difficult to score and print out new Accelerated Math assignments. In some cases, students found themselves with no set task to work on, causing the supervisors to think quickly and find a new task from a booklet or problem sets in the textbook.

Another logistical limitation with the study was the dual role of the researcher/classroom teacher. While the academic data was likely purely objective and unbiased, the survey data does risk bias since some students may not trust that they were not being judged by their responses, that their opinions would not affect their grades or how the classroom teacher thought of them. As the researcher and classroom teacher, I did my best to reiterate to the students that their responses would not in any way be used against them, that their survey choices—positive or negative—would have no bearing on their interactions with me, their classroom teacher, nor were their answers being scrutinized or judged in any way. The survey was presented as a means to collect formative data, the exact purpose for which it was designed. There is no guarantee that the survey data is biased or unbiased, but the possibility certainly needs to be acknowledged.

CHAPTER 5: Conclusion

The purpose of this study was to determine whether a combination of interventions--peer tutoring and additional learning time--increases students' attitudes about and achievement in mathematics. As noted in the Discussion section, the results are inconclusive. While there was no clearly and consistently significant increase in posttest scores in some of the groups, there were notable increases in some areas such as some of the survey categories. Regardless, there was not a significant decrease in the treatment results either, a phenomenon that could have happened without the treatment.

This study added to the current body of research regarding peer tutoring and out-of-school time. There was a need for more quantitative analysis that utilizes control and treatment groups in the area of peer tutoring, specifically regarding students in middle school, and this study addressed that need. This study also contributed to the out-of-school time research by providing a structured environment with clear expectations by utilizing the Accelerated Math program as a supplement, though there were technological obstacles. OST research also was lacking in information regarding students in middle school, which this study has addressed directly. Much has been learned throughout this study, and ground has been broken for future research.

With more time and resources, there are many more questions this study could answer. Most of the students in the treatment group had lower pretest scores than their peers in the control group, and it could be concluded that, based on the nearly identical percentage point increases between the pre- and posttests, the treatment group made the same degree of progress as that of the control group. Therefore, the gap between the two groups that had been observed at the beginning of the year did not grow. Because this study saw no evidence of that discrepancy increasing, it remains to be seen whether the interventions were the direct cause of the commensurate increases in both posttest and postsurvey results. A longitudinal study of control groups and treatment groups such as these would help answer that question. Also, collecting data for a second control group of students with characteristics similar to those in the treatment group could help determine whether the interventions were effective.

References

Aidman, B. J. (1997). How Summer School Can Provide a Jump Start for Students. Educational Leadership, 55, 64-68.

Baker, J., Rieg, S., & Clendaniel, T. (2006). An Investigation of an After School Math Tutoring Program: University Tutors + Elementary Students = Successful Partnership. Education, 127 (2), 287-293.

Barone, M. & Taylor, L. (1996). Peer Tutoring. Teaching Children Mathematics, 3, 8-15

Dufrene, B., Noell, G., & Gilbertson, D. (2005). Monitoring Implementation of Reciprocal Peer Tutoring: Identifying and Intervening With Students Who Do Not Maintain Accurate Implementation. The School Psychology Review, 34 (1), 74-86.

Fennema, E. H. & Sherman, J. A. (1976). Fennema-Sherman Mathematics Attitudes Scales: Instruments designed to measure attitudes toward the learning of mathematics by males and females. Catalog of Selected Documents in Psychology, 6(1), 31.

Fuchs, L., Compton, D., Fuchs, D., Paulsen, K., Bryant, J., & Hamlett, C. (2005). The Prevention, Identification, and Cognitive Determinants of Math Difficulty. Journal of Educational Psychology, 97 (3), 493-513.

Gaeddert, T. (2001). Using Accelerated Math to Enhance Student Achievement in High School Mathematics Courses. Renaissance Learning, Inc.

Ginsburg-Block, M. D., Rohrbeck, C. A., & Fantuzzo, J. W. (2006). A meta-analytic review of social, self-concept, and behavioral outcomes of peer-assisted learning.

Journal of Educational Psychology, 98 (4), 732-749.

Jekielek, S M. Moore, K. A.; Hair, E. C.; Scarupa, H. J. (2002). Mentoring: A promising strategy for youth development. Child trends research brief. Retrieved March 1, 2010 from http://www.childtrends.org..

Kroesenbergen, E. & Van Luit, J. (2003). Mathematics Interventions for Children with Special Educational Needs. Remedial & Special Education, 24 (2), 97-104.

Lauer, P., Akiba, M., Wilkerson, S., Apthorp, H., Snow, D., & Martin-Glenn, M. (2006). Out-of-School-Time Programs: A Meta-Analysis of Effects for At-Risk Students. Review of Educational Research, 76 (2), 275-313.

Miller, B.M. (2003). Critical Hours: Afterschool Programs and Educational Success. Quincy, MA: Nellie Mae Education Foundation. Retrieved February 24, 2010, from http://www.nmefdn.org/uploads/Critical_Hours.pdf.

Mulhern, F. & Rae, G. (1998). Development of a Shortened Form of the Fennema-Sherman Mathematics Attitudes Scales. Educational and Psychological Measurement, 58 (2), 295-306.

O'Neal, M., Ernest, P., McLean, J., & Templeton, S. M. (1988). Factorial Validity of the Fennema-Sherman Mathematics Attitudes Scales. Paper presented at the Annual Meeting of the Mid-South Educational Research Association. Louisville, Kentucky, November 9-11, 1988.

Robinson, D.R., Schofield, J.W., & Steers-Wentzell, K.L. (2005). Peer and Cross-Age Tutoring in Math:Outcomes and Their Design Implications. Educational Psychology Review, 17 (4), 327-362.

Sandusky, L. & Brem, S. (2002). The Use of Accelerated Math in an Urban Title I Elementary School. 2001-2002 Supplementary Report to Renaissance Learning, Inc., Arizona State University.

Schoenfeld, A.H. (1989). Explorations of Students' Mathematical Beliefs and Behavior. Journal for Research in Mathematics Education, 20 (4), 338-355.

Spicuzza, R., Ysseldyke, J., Lemkuil, A., McGill, S., Boys, C., & Teelucksingh, E. (2001). Effects of Curriculum-Based Monitoring on Classroom Instruction and Math Achievement. Journal of School Psychology, 39 (6).

Springer, R., Pugalee, D., Algozzine, B. (2007). Improving Mathematics Skills of High School Students. The Clearing House, 81 (1), 37-44.

Steele, D. & Arth, A. (1998). Math Instruction and Assessment: Preventing Anxiety, Promoting Confidence. Schools in the Middle, 6, 44-48.

Topping, K., Campbell, J., Douglas, W., & Smith, A. (2003). Cross-age peer tutoring in mathematics with seven- and 11-year-olds: influence on mathematical vocabulary, strategic dialogue and self-concept. Educational Research, 45 (3), 287-308.

Walker, E. (2007). The Structure and Culture of Developing a Mathematics Tutoring Collaborative in an Urban High School. The High School Journal, 91 (1), 57-67.

Ysseldyke, J., Betts, J., & Hannigan, E. (2004). Use of an Instructional Management System to Improve Mathematics Skills for Students in Title I Programs. Preventing School Failure, 48 (4), 10-14.

Ysseldyke, J., Spicuzza, R., Kosciolek, S., Boys, C. (2003). Effects of a Learning Information System on Mathematics Achievement and Classroom Structure. Journal of Educational Research, 96 (3), 163-173.

Ysseldyke, J. & Tardrew, S. (2002). Differentiating Math Instruction: A Large Scale Study of Accelerated Math. Renaissance Learning, Inc.

Zuelke, D. & Nelson, J. G. (2001). The Effects of a Community Agency's After-School Tutoring Program on Reading and Math GPA for At-Risk Tutored Students. Education, 121 (4), 799-899.

Appendix

Table 14
Correlation between Students' Academic Pretests and Posttests Controlling for Services

Control Variables			Pretest	Posttest
Services	Pretest	Correlation	1.000	.638
		Significance (2-tailed)	.	.000
		df	0	96
	Posttest	Correlation	.638	1.000
		Significance (2-tailed)	.000	.
		df	96	0

Table 15
Correlation between Students' Academic Pretests and Posttests Controlling for Hours Tutored

Control Variables			Pretest	Posttest
Hrs	Pretest	Correlation	1.000	.671
		Significance (2-tailed)	.	.000
		df	0	96
	Posttest	Correlation	.671	1.000
		Significance (2-tailed)	.000	.
		df	96	0

Table 16
Correlation between Students' Academic Pretests and Posttests Controlling for Hours Tutored

		Services	Percentage Point Difference
Services	Pearson Correlation	1	-.078
	Sig. (2-tailed)		.445
	N	112	99
Percentage Point Difference	Pearson Correlation	-.078	1
	Sig. (2-tailed)	.445	
	N	99	99

Table 17
Comparison of Attitudinal Survey Means by Category

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% CI	
		F	Sig.	t	df	Sig. (2-tailed)	*MD*	*SE*	*LL*	*UL*
Confidence-Change	Equal variances assumed	6.07	.02	.82	101	.42	1.01	1.24	-1.45	3.48
	Equal variances not assumed			.99	87.52	.32	1.01	1.02	-1.01	3.03
Usefulness-Change	Equal variances assumed	4.83	.03	.66	101	.51	.77	1.17	-1.54	3.09
	Equal variances not assumed			.52	36.11	.60	.77	1.48	-2.23	3.78
Male Domain-Change	Equal variances assumed	4.78	.031	-1.72	101	.09	-1.315	.77	-2.834	.20
	Equal variances not assumed			-1.25	32.81	.22	-1.32	1.05	-3.46	.83
Success-Change	Equal variances assumed	.40	.53	1.14	101	.26	1.00	.88	-.74	2.74
	Equal variances not assumed			1.01	42.97	.32	1.00	.99	-1.00	2.99
Teacher-Change	Equal variances assumed	1.31	.26	.22	101	.83	.18	.84	-1.48	1.85
	Equal variances not assumed			.21	50.61	.83	.18	.87	-1.56	1.92
Anxiety-Change	Equal variances assumed	.001	.98	1.75	101	.08	1.49	.85	-.20	3.18
	Equal variances not assumed			1.76	54.41	.08	1.49	.85	-.21	3.19
Effective Motivator-	Equal variances assumed	.01	.91	.39	101	.70	.12	.31	-.49	.72

Change	Equal variances not assumed			.38	52.05	.70	.12	.31	-.50	.74
Schoenfeld-Change	Equal variances assumed	8.19	.01	-.29	101	.77	-.28	.97	-2.19	1.64
	Equal variances not assumed			-.24	39.71	.81	-.28	1.14	-2.59	2.04

Table 18
Comparison of Means for Attitudinal Survey Categories: Math 7

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	*MD*	*SE*	*LL*	*UL*
Confidence-Change	Equal variances assumed	3.08	.09	-.35	39	.73	-.66	1.88	-4.47	3.15
	Equal variances not assumed			-.44	35.23	.66	-.66	1.50	-3.71	2.39
Usefulness-Change	Equal variances assumed	2.61	.11	.27	39	.79	.60	2.22	-3.90	5.10
	Equal variances not assumed			.20	12.66	.85	.60	3.03	-5.96	7.15
Male Domain-Change	Equal variances assumed	9.07	.01	-1.82	39	.08	-2.91	1.60	-6.16	.33
	Equal variances not assumed			-1.19	11.34	.26	-2.91	2.45	-8.28	2.45
Success-Change	Equal variances assumed	1.66	.21	.12	39	.91	.19	1.62	-3.10	3.48
	Equal variances not assumed			.09	14.01	.93	.19	2.03	-4.17	4.55
Teacher-Change	Equal variances assumed	.03	.86	-.14	39	.89	-.22	1.65	-3.56	3.11
	Equal variances not assumed			-.14	21.20	.89	-.22	1.63	-3.61	3.16

Anxiety-Change	Equal variances assumed	.47	.50	1.74	39	.09	2.47	1.42	-.41	5.34
	Equal variances not assumed			1.98	28.16	.06	2.47	1.25	-.09	5.02
Effectance Motivation-Change	Equal variances assumed	.80	.38	-.36	39	.72	-.18	.51	-1.21	.85
	Equal variances not assumed			-.34	17.79	.74	-.18	.55	-1.34	.97
Schoenfeld-Change	Equal variances assumed	8.53	.01	-1.05	39	.30	-1.66	1.58	-4.86	1.53
	Equal variances not assumed			-.78	12.75	.45	-1.66	2.14	-6.29	2.96

Table 19
Comparison of Means for Attitudinal Survey Categories: Math 8

		Levene's Test		t-test for Equality of Means						
									95% CI	
		F	Sig.	t	df	Sig. [a]	*MD*	*SE*	*LL*	*UL*
Confidence-Change	Equal variances assumed	5.12	.05	1.81	11	.10	3.17	1.75	-.68	7.01
	Equal variances not assumed			1.30	3.40	.28	3.17	2.44	-4.11	10.44
Usefulness-Change	Equal variances assumed	.07	.80	.14	11	.90	.42	3.09	-6.38	7.22
	Equal variances not assumed			.15	7.78	.88	.42	2.75	-5.96	6.79
Male Domain-Change	Equal variances assumed	.01	.94	.33	11	.75	.31	.94	-1.77	2.38
	Equal variances not assumed			.33	6.17	.75	.31	.92	-1.93	2.54
Success-Change	Equal variances assumed	.86	.38	.64	11	.53	1.67	2.59	-4.03	7.36
	Equal variances not assumed			.58	4.63	.59	1.67	2.90	-5.96	9.29

Teacher-Change	Equal variances assumed	1.57	.24	.55	11	.59	1.03	1.87	-3.08	5.14
	Equal variances not assumed			.69	10.13	.51	1.03	1.49	-2.29	4.34
Anxiety-Change	Equal variances assumed	.000	.98	-.79	11	.45	-2.42	3.08	-9.19	4.36
	Equal variances not assumed			-.76	5.47	.48	-2.42	3.17	-10.35	5.52
Effectance Motivation-Change	Equal variances assumed	.98	.34	.91	11	.38	.89	.97	-1.26	3.03
	Equal variances not assumed			.76	4.11	.49	.89	1.17	-2.33	4.11
Schoenfeld-Change	Equal variances assumed	.04	.84	.08	11	.94	.28	3.35	-7.10	7.66
	Equal variances not assumed			.08	4.70	.94	.28	3.72	-9.47	10.03

a. 2-tailed

Table 20
Comparison of Means for Attitudinal Survey Categories: Pre-Algebra

		Levene's Test		t-test for Equality of Means						
									95% CI	
		F	Sig.	t	df	Sig. [a]	*MD*	*SE*	*LL*	*UL*
Confidence-Change	Equal variances assumed	2.49	.12	.46	47	.65	1.08	2.37	-3.69	5.85
	Equal variances not assumed			.59	17.53	.56	1.08	1.83	-2.77	4.93
Usefulness-Change	Equal variances assumed	4.20	.05	.77	47	.44	1.29	1.67	-2.07	4.66
	Equal variances not assumed			.58	9.52	.57	1.29	2.22	-3.68	6.27
Male Domain-Change	Equal variances assumed	.03	.85	-.82	47	.41	-.81	.98	-2.77	1.16

	Equal variances not assumed			-.82	11.87	.43	-.81	.98	-2.94	1.33
Success-Change	Equal variances assumed	.21	.65	1.47	47	.15	1.87	1.27	-.68	4.42
	Equal variances not assumed			1.57	12.78	.14	1.87	1.19	-.71	4.45
Teacher-Change	Equal variances assumed	1.77	.19	.97	47	.34	1.14	1.18	-1.22	3.51
	Equal variances not assumed			.80	10.05	.44	1.14	1.43	-2.03	4.32
Anxiety-Change	Equal variances assumed	.05	.82	1.41	47	.17	1.77	1.25	-.75	4.29
	Equal variances not assumed			1.37	11.50	.20	1.77	1.29	-1.06	4.60
Effectance Motivation-Change	Equal variances assumed	1.76	.19	1.39	47	.17	.62	.44	-.28	1.51
	Equal variances not assumed			1.76	16.66	.10	.62	.35	-.12	1.36
Schoenfeld-Change	Equal variances assumed	1.29	.26	.82	47	.42	1.23	1.50	-1.80	4.25
	Equal variances not assumed			.71	10.45	.49	1.23	1.73	-2.61	5.06

a. 2-tailed

Table 21
Effect Sizes for Means for Each Attitudinal Survey Category by Class

Class	Category	Cohen's *d*	Effect Size *r*
Math 7	Confidence-Change	-.13	-.07
	Usefulness-Change	.08	.04
	Male Domain-Change	-.48	-.23
	Success-Change	.04	.02
	Teacher-Change	-.05	-.02
	Anxiety-Change	.64	.30
	Effectance Motivation-Change	-.12	-.06
	Schoenfeld-Change	-.30	-.15
Math 8	Confidence-Change	.89	.40
	Usefulness-Change	.09	.04
	Male Domain-Change	.20	.10
	Success-Change	.36	.18
	Teacher-Change	.37	.18
	Anxiety-Change	-.46	-.23
	Effectance Motivation-Change	.49	.24
	Schoenfeld-Change	.05	.02
Pre-Algebra	Confidence-Change	.19	.09
	Usefulness-Change	.24	.12
	Male Domain-Change	-.31	-.15
	Success-Change	.52	.25
	Teacher-Change	.32	.16
	Anxiety-Change	.51	.25
	Effectance Motivation-Change	.57	.28
	Schoenfeld-Change	.28	.14