

Micki M. Caskey, Ph.D., Editor
Portland State University
Portland, Oregon

2011 • Volume 34 • Number 6

ISSN 1940-4476

Do School-Based Tutoring Programs Significantly Improve Student Performance on Standardized Tests?

Terri Rothman
Monmouth University
West Long Branch, NJ

Mary Henderson
Long Branch High School
Long Branch, NJ

Abstract

This study used a pre-post, nonequivalent control group design to examine the impact of an in-district, after-school tutoring program on eighth grade students' standardized test scores in language arts and mathematics. Students who had scored in the *near-passing* range on either the language arts or mathematics aspect of a standardized test at the end of seventh grade were recruited to receive tutoring in either language arts (LA) or mathematics (MA), depending on the area of weakness. An analysis of covariance revealed that both groups of students tutored in LA ($n = 23$) or MA ($n = 20$) significantly outperformed a matched control group ($p = .02$ for LA; $p = .04$ for MA). Components of effective after-school academic programs are discussed.

Background

In the United States, the No Child Left Behind Act mandates that schools ensure all students pass state proficiency exams in reading and math by the 2013–2014 academic year. Schools receiving Title I funds that have not made adequate yearly progress for three or more years are required to offer parents an opportunity for their children to receive supplemental educational services (SES), usually in the form of tutoring. Although this provision has been part of the law since its inception, Borja (2007) reported that only 23% of eligible students (585,000 of 2.5 million) received NCLB tutoring in 2005–2006.

Not a single state has provided data to show that students enrolled in these programs have made appreciable gains (Lewis, 2006). In addition, the 2006 U.S. Department of Education's Title I report indicated that the number of Title I schools identified as "in need of improvement" for failing to make adequate yearly progress nearly doubled between 2005 and 2006 (Davis, 2006).

Although little evidence supports the effectiveness of SES tutoring programs, tutoring has a long documented history as a reliable method to improve student achievement (Slavin, 1999). Tutoring, as a supplement to classroom teaching, is generally considered the most powerful form of instruction for increasing underachieving students' reading achievement (Burns, Senesac, & Symington, 2004). The U.S. Department of Education remains committed to tutoring as a remedy for student underachievement and is gathering data from pilot programs in Boston, Chicago, and New York. These programs permit districts to use their own teachers in school-based tutoring programs scheduled at the end of the school day.

Much literature indicates the positive impact of quality relationships on student success in the middle grades. The Turning Points approach suggests, "Reform is most successful when schools receive intensive, on-site support from experienced educators who reside close to the schools they serve" (Feldman & Ouimette, 2004, p. 2). Under this approach, site-based teachers serve as in-house facilitators in the whole-school reform process (Feldman & Ouimette). A report by Cassellius (2006) showed recent success in the Memphis City schools, with middle school students showing increased standardized test performance in both math and language arts. She stated that the key to school success is the relationships between people (Cassellius); in addition, high expectations of all students within a safe learning environment are also critical (Cassellius). In a recent report on effective middle school reform practices, Jackson (2009) described the need for teacher development, school-community relationships, and the need to focus on global perspectives. With the increasing diversity of the student population, knowledge of different cultures is necessary to promote inter-cultural relationships (Jackson).

The current study seeks to extend the body of research on school-based tutoring programs that use district teachers as tutors. Given the paucity of evidence to substantiate the effectiveness of SES tutoring programs, the current study examined differences in performance on standardized tests in the areas of language arts and mathematics between participants and nonparticipants in an after-school tutoring program.

Effective After-School Tutoring Programs

A comprehensive report (Fashola, 1998) on effective after-school programs investigated formal and informal programs that varied by location (school-based vs. community-based) and type of provider

(classroom teachers vs. childcare providers). Results demonstrated that students who attended a formal program that was housed in the students' school and staffed by classroom teachers outperformed the other groups in mathematics and reading. Fashola recommended, "Content taught during the after-school period must be taught by qualified instructors who are familiar with and can be held accountable for student outcomes" (p. 49). Furthermore, he suggested staffing after-school programs with regular-school-day teachers as an efficient method to ensure alignment of the after-school curriculum with the school curriculum. Gordon (2009) agreed that the most effective tutoring programs use master teachers who continuously collaborate with the students' classroom teachers.

A recent meta-analytic report (Lauer, Akiba, Wilkerson, Apthorp, Snow, & Martin-Glenn, 2006) investigated and summarized research reports and evaluations of out-of-school-time (OST) programs for students at risk for school failure; the meta-analysis included programs that focused on reading, math, or both. Researchers specifically analyzed the effects of time frame (after school, summer, Saturday), grade level, program focus (academic, academic + social), duration, and instructional grouping. Results suggested that OST programs positively affected the reading and math achievement of students at risk for school failure, whether programs were offered after school, during the summer, or on Saturdays. Programs of moderate duration (45–85 hours) had the greatest impact on both reading and math achievement. Further, programs that focused on both academics and social aspects were as effective as purely academic programs.

Slight differences emerged in Lauer and associates' (2006) study, however, when comparing math and reading programs. For example, reading programs benefited both elementary and secondary students, whereas benefits in mathematics only pertained to secondary students. In addition, one-on-one tutoring had the strongest positive effect on reading achievement, while small and mixed instructional grouping practices had the greatest impact on mathematics achievement.

A recent meta-analysis (Ritter, Barnett, Denny, & Albin, 2009) attempted to identify components of effective volunteer programs by investigating studies of students in grades K–8 with randomized field trials. They found that volunteer tutors had a positive impact on student performance in reading, but little

evidence was available for math. Furthermore, type of volunteer (e.g., college student, teacher, parent) did not have a differential impact on student outcomes.

Caring tutoring relationships. The tutor-tutee relationship appears to be important to program success; when students perceive that a teacher cares about his or her students, academic achievement improves (Klem & Connell, 2004). District teachers who have established personal relationships with students have an advantage over new, unfamiliar tutors. Students often regard district teachers as “institutional representatives,” signaling to students that not only does this teacher care for them but also that they are worth the effort and belong in school (Black, 2003). Positive perceptions of teacher caring also led to higher levels of student engagement (Skinner & Belmont, 1993; Triplett, 2004), which are, in turn, associated with higher attendance rates and higher test scores (Klem & Connell, 2004). When out-of-district tutors are hired, they lack this familiarity with the children and lack any connection with the foundational school or a child’s general sense of belonging.

Collaboration with the classroom teacher is often cited as vital to a tutoring program’s success or failure (Gordon, 2009; Gordon, Morgan, Ponticell, & O’Malley, 2004). District tutors have daily contact with tutees’ subject area teachers, thus facilitating communication about the classroom curriculum and the impact of tutoring on classroom performance. This dialog between subject area teacher and tutor provides a reciprocal benefit. The teacher tutor helps improve instruction for the student in the regular classroom, and the regular classroom teacher can convey important information to the teacher tutor, which enhances the tutoring sessions. As the tutoring partnership progresses, the student begins to see a clearer connection between tutoring and school work. This connection to the school setting is vital to the student’s academic improvement and is more easily achieved by a district teacher.

A review of 27 studies of after-school programs demonstrated a relationship between attendance rates and positive student outcomes (McComb & Scott-Little, 2003), yet program evaluation reports seldom present attendance rates. Moreover, inconsistent attendance threatens the ability to determine a program’s impact on student outcomes (Hock, Pulvers, Deshler, & Schumaker, 2001). School-based teachers are in a unique position to influence attendance positively. For example, the school-based teacher has greater access to parents, who can monitor and encourage their children’s attendance. In addition, these

school-based tutors can communicate with a student’s final period teacher or visit the final period to remind the student about attending the program. Ultimately, using a school-based tutor lends itself more easily to a teamwork approach to helping the child improve.

Much of the research suggested that longer partnerships with students provide greater and longer-lasting results (Baker, Reig, & Clendaniel, 2006). Researchers suggested a minimum of 30 hours of tutoring is needed to show student improvement (Gordon et al., 2004), with some opportunities for one-on-one tutoring (Fashola, 1998). Despite the relatively consistent recommendations on the minimum program duration, other reports provide contrasting viewpoints on the upper limits of program duration. Pascopella (2004) reported that students receiving 160 hours of tutoring showed a six-month gain in reading and a two-year gain in math. In contrast, Lauer and associates’ (2006) meta-analysis indicated that programs of 160 hours’ duration positively affected reading performance but had a negligible effect on mathematics performance. Regardless of the time allocation, using district teachers and facilities offered greater opportunity to ensure that students attend the program for the required number of hours (Fashola, 1998).

In an attempt to add needed data to this field of research, this study examined the impact of an after-school program using district teacher tutors on student performance on standardized tests. Standardized test scores of eighth grade students who were tutored in either language arts or math were compared to scores of a matched control group that did not receive tutoring. It was hypothesized that the teacher-tutored groups would score higher in language arts (LA) and math (MA) than the non-tutored students.

Method

Setting and Sample

This study was conducted in one large, ethnically diverse, urban central New Jersey public school district consisting of three elementary schools, one middle school (grades 6–8), and one high school (grades 9–12). The district is classified as an Abbott district, which means it meets specific criteria, including (1) low socioeconomic status, according to the New Jersey Department of Education; (2) evidence of substantive failure of thorough and efficient education; (3) a large percentage of disadvantaged students who need “an education beyond the norm;” and (4) the existence of an “excessive tax for municipal services.”

This study was conducted at the middle school level, which is housed within one school. The middle school student population consists of 295 sixth, 321 seventh, and 334 eighth graders. Of these students, 60% are eligible for free or reduced-price lunch, with a nearly even breakdown of three major ethnic groups (35% Latino, 33% African American, and 32% Caucasian). Approximately 30% are non-native English speakers (Spanish or Portuguese), although only five percent meet the definition of limited English proficient.

The targeted population for this study consisted of 102 eighth grade students (30% of the eighth grade) who were designated as “borderline” based on scoring 180–199 on the seventh grade New Jersey Assessment of Skills and Knowledge (NJASK). Of the 102 students who met these criteria, all were purposefully selected to participate in one of the treatment groups, and 43 agreed to participate. Participants were specifically assigned to one of two treatment groups, LA or MA, depending on which area was borderline. In cases where a student scored borderline on both LA and MA, the student was randomly assigned to only one tutoring group. The treatment (tutoring) groups consisted of 43 volunteers (LA = 23; MA = 20); 12 were Caucasian, 13 were Latino, and 18 were African American students. Ages ranged from 13 to 15.

Of the remaining 59 non-volunteers, 37 participants (17 females, 20 males, ages 13 to 15) were randomly selected to serve as the control group. The ethnic breakdown was as follows: 8 Caucasian, 15 Latino, and 14 African American students.

Tutors and training. Tutor training is an essential component of any effective after-school program (Fashola, 1998) and is directly related to fidelity of implementation. As noted in a recent review of the literature (O’Donnell, 2008), fidelity of implementation refers to adherence to protocol and maintaining the overall integrity of the program. Following protocol is important to determine whether intended outcomes are the result of the program under evaluation. However, O’Donnell noted the extensive divergence in both defining and evaluating fidelity. For the purpose of this study, the researchers attempted to improve fidelity through regular tutor meetings with discussions of the content of the tutoring sessions and potential problems detected during the sessions. Because tutors followed the program design as intended, outcomes may be attributed to the implementation of the program, which provides evidence of fidelity (O’Donnell); however, if the desired outcomes do not occur, the program will need modification. To increase confidence in fidelity of

implementation, the researchers created a checklist to be used in future tutoring programs.

Eleven district tutors, who volunteered to participate in the tutoring program, were selected based on perceived effectiveness in the classroom. For example, their students consistently demonstrated higher test scores than students in similar classes with other teachers. These teachers rarely referred their students to the office for disciplinary reasons, thus their student-teacher relationships were also perceived as strong. The strongest teachers were selected for the program, with the philosophy that the program would primarily be a continuation of the school day. The tutoring setting (i.e., classrooms within the middle school) provided a much smaller teacher-student ratio, providing more intimate contact and individualized instruction.

Because previous research showed that individualized and small-group instruction enhanced the positive effects of a summer program for both low-income and middle-income students (Cooper, Charlton, Valentine, & Muhlenbruck, 2000), the researchers elected to use small-group instruction. The student-teacher ratio was 4:1; each teacher tutored one group of four students. Tutoring materials included *Preparing for the New Jersey Grade Eight Proficiency Assessment* (GEPA) (2001) booklets and Standard Solutions (2006) test-taking strategies worksheets.

Students attended tutoring sessions two afternoons per week for 90 minutes each day directly following the end of the regular school day from October through March, for a total of 48 hours, which is sufficient program time to allow an effective evaluation (Slavin, 2008). Tutors provided rewards including pizza parties and bowling parties for perfect attendance.

Both math and language arts teacher-tutors taught from the *Preparing for the New Jersey GEPA* booklet (2001), a different curriculum than that taught during the school day. Equally important, teachers were trained to use the Standard Solutions (2006) test-taking strategies while working within the curriculum. The Standard Solutions program teaches students strategies that increase their test-taking abilities. Strategies include identification and elimination of multiple-choice options and proper completion of all components of open-ended questions. Effective time management strategies during testing sessions are also addressed.

Teacher-tutors were required to attend four meetings during the program. They also provided weekly progress reports to classroom teachers. Their records

revealed a 90 % average student attendance rate for the after-school tutoring program.

Instruments

The New Jersey Assessment of Skills and Knowledge (NJASK), a standardized state exam, was administered for the first time in April 2006 to students in seventh grade in New Jersey. The NJASK consists of two content areas: mathematics (numbers, numerical operations, geometry and measurement, patterns and algebra, data analysis, probability, and discrete mathematics) and language arts literacy (writing, speaking, reading, listening, viewing, working with or interpreting text, analyzing/critiquing text, and extending understanding of text). The exam consists of both closed-ended and open-ended questions. The NJASK was developed by the Educational Testing Service (ETS) and is generally accepted as having strong content validity, as it is aligned with the New Jersey Core Curriculum Content Standards. NJASK language arts and math scores in seventh grade were used as covariates in this study.

Language arts and math scores (consisting of content topics that parallel those found on the NJASK) on the Grade Eight Proficiency Assessment (GEPA) were used as the dependent variables. The GEPA, a standardized state exam developed by ETS, is viewed as having strong content validity and reliability.

Research Design

The design used in this study was a pre-post, nonequivalent control group design (see Table 1). There were two treatment groups: one group received tutoring in language arts, and the second group received tutoring in math.

Procedure

Prior to the beginning of the 2006–2007 academic year, 102 students were selected as a target group to receive tutoring. Using the list of seventh grade NJASK standardized test scores, students who scored between

180 and 199 (“borderline”) were selected to attend the program. These students were purposefully selected because a 20-point gain was deemed remediable under the parameters of the program. The researchers believed that the additional support of the tutoring program would adequately enable borderline students to make the gains necessary to pass the eighth grade test. Students scoring below 180 were not selected because they might have needed additional instruction beyond what could be provided within the October to March time frame. Invitations were distributed to all 102 students who met the criteria, and 43 students agreed to attend the 48-hour program. From the remaining 59 students within the same score range, 37 were randomly selected for the control group.

Treatment groups. Students were placed in one of two treatment (tutoring) groups, depending on which NJASK score (language arts or math) was borderline. Twenty-three students were tutored in language arts, and 20 were tutored in math. Students were told that they were selected because they were the group most likely to pass if given additional assistance. This was a consistent message given to the students, and teachers were told during tutor meetings to reinforce this idea with the children.

At the end of the 48-hour program, two analyses were conducted. Using verbal scores on the NJASK as the covariate, students who were tutored in language arts were compared to the control group on GEPA verbal scores. Using math scores on the NJASK as the covariate, students who were tutored in math were compared to the control group on GEPA math scores. Covariates were selected to control for initial differences between the experimental and control groups. The NJASK verbal and math scores were deemed appropriate measures of initial group differences on verbal or math performance, as the NJASK is a standardized instrument that was administered at the end of seventh grade.

Table 1
Pre-Post Nonequivalent Control Group Design: Treatment Consists of Language Arts (LA) Tutoring or Math (MA) Tutoring

Group	Assignment	n	Treatment	Pretest	Posttest
LA Tutoring	Volunteer	23	Tutoring-LA	NJASK(LA)	GEPA (LA)
Matched Control	Random	37	No tutoring	NJASK(LA)	GEPA (LA)
Math Tutoring	Volunteer	20	Tutoring-MA	NJASK(MA)	GEPA (MA)
Matched Control	Random	37	No tutoring	NJASK(MA)	GEPA (MA)

Data Analysis

Because the treatment group consisted of volunteers from the population of students at risk for school failure, an Analysis of Covariance (ANCOVA) was used. Scores on the NJASK, administered at the end of seventh grade, were used as covariates, enabling the researchers to equate the groups for any initial differences in performance on standardized tests. This type of analysis provided the means to determine whether differences between the treatment and control groups on performance on the GEPA administered at the end of eighth grade were due to the tutoring program. Data analyses indicated that the assumptions (e.g., homoscedasticity, sphericity) for conducting an ANCOVA were met.

Results

This study compared scores on standardized tests in either mathematics or language arts between each of two treatment groups (i.e., tutoring in language arts

or mathematics) and a control group (i.e., no tutoring). The pretest and posttest mean scores for both the group that received language arts tutoring and the control group are displayed in Table 2. To determine if tutoring had an impact on students who were at risk of performing poorly on the GEPA exam, the researchers calculated an ANCOVA. Using language arts scores on the NJASK as a covariate, the results indicated that students who were tutored in language arts significantly outperformed the control group on the language arts section of the GEPA ($F(1, 57) = 5.835, p = .02$), which are shown in Table 3.

The pretest and posttest mean scores for both the group that received math tutoring and the control group are presented in Table 4. Using math scores on the NJASK as a covariate, the results indicated that students who were tutored in math outperformed the control group on the math section of the GEPA ($F(1, 54) = 4.55, p = .04$), as displayed in Table 5.

Table 2
Pretest and Posttest Language Arts Means and Standard Deviations

Group	Sample (n)	Ethnicity*			Pretest means		Posttest means	
					NJASK (LA)		GEPA (Verbal)	
					Mean	SD	Mean	SD
LA Tutoring	23	13	4	6	201.08	19.99	203.17	27.39
No tutoring	37	13	6	18	198.68	18.85	188.08	29.03

Note. Sample and ethnicity numbers represent actual number of participants in each category.
* A= African American; C = Caucasian; L = Latino

Table 3
Tests of Between-Subjects Effects for the Language Arts Group Dependent Variable: Language Arts Scores on the GEPA

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	30916.91(a)	2	15458.45	46.00	.00
Intercept	477.18	1	477.18	1.42	.24
NJASK Language Arts	27686.04	1	27686.04	82.39	.00
Group	1960.70	1	1960.70	5.84	.02
Error	19154.03	57	336.04		
Total	2305128.00	60			
Corrected Total	50070.933	59			

Note. R Squared = .62 (Adjusted R Squared = .60)

Table 4
Pretest and Posttest Mean Math Scores and Standard Deviations

Group	Sample (n)	Ethnicity*			Pretest means		Posttest means	
		A	C	L	NJASK (MA)		GEPA (Math)	
					Mean	SD	Mean	SD
Math Tutoring	20	5	8	7	188.90	10.42	197.05	25.57
No tutoring	37	13	6	18	191.86	18.76	188.76	25.09

Note. Sample and ethnicity numbers represent actual number of participants in each category.

* A= African American; C = Caucasian; L = Latino

Table 5
Tests of Between-Subjects Effects for the Math-Tutored Group Dependent Variable: Math Scores on the GEPA

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	13699.92	2	6849.96	19.79	.00
Intercept	95.59	1	95.59	.28	.60
NJASK Math	12807.01	1	12807.01	37.01	.00
Group	1575.17	1	1575.17	4.55	.04
Error	18688.75	54	346.09		
Total	2126347.00	57			
Corrected Total	32388.67	56			

Note. R Squared = .42 (Adjusted R Squared = .40)

Discussion

The results of this study indicate that borderline students who received school-based tutoring from district teachers performed higher on standardized test scores in the areas of mathematics and language arts than borderline students who did not participate in tutoring. These findings contribute to the growing body of literature on out-of-school tutoring programs (Ritter et al., 2009).

As part of the No Child Left Behind Act, schools receiving Title I funds that have not made adequate yearly progress for three or more years are required to offer students SES, often in the form of tutoring. While this provision was originally intended for outside agencies, little data revealed appreciable gains made by students enrolled in these programs (Borja, 2007). A recent meta-analysis (Ritter et al., 2009) demonstrated the lack of empirical evidence for programs that focused on math, although out-of-district tutoring programs appeared to impact reading performance. Underutilization of the contracted tutoring services, coupled with a lack of documentation of success, led the New Jersey

State Department of Education to allow a number of districts to provide the services. The current study supports the policy change that would allow district teachers to provide tutoring to students to improve standardized test scores in both reading and math.

Although some prior research failed to demonstrate positive effects of tutoring on standardized test performance, the use of district tutors may have been a contributing factor in the current study, as tutor knowledge of the curriculum and tutor-student relationships are important components of a successful tutoring program. For example, Morris (2005) found that using the school’s teachers and paraprofessionals was very effective in improving reading ability. Furthermore, Zuelke and Nelson (2001) found little impact of tutoring programs, even when teachers were the tutors, if tutors failed to communicate effectively with the child’s classroom teacher. Perhaps district teachers who have an intimate knowledge of student and curriculum may be better equipped than contracted outsiders to provide effective tutoring instruction to students. Further research is warranted to compare these two different types of tutors.

Lauer and associates' (2006) meta-analytic study provided crucial information about effective components of OSP programs, and several aspects of the current tutoring program using district teachers correspond with these essential elements. First, the program was 48 hours in length, which is within the medium-length time range that Lauer and associates reported to be the most effective for both reading and mathematics performance; longer programs (100 hours for math; 210 for reading) produce smaller effects. Second, the program incorporated pizza parties and bowling parties; however, according to Lauer and associates the addition of a social aspect does not preclude success. Also included in the current program was small-group instruction, which has been found to be beneficial, especially in the area of mathematics. For reading, one-on-one tutoring may be the most effective format, according to Lauer and associates.

Although the literature points to a variety of specific program conditions that appear to contribute to the success or failure of tutoring programs, the student-teacher relationship is consistently identified as a key factor for increasing student achievement (Slavin, 1999). When students perceive that teachers care about their well-being, school engagement increases, which ultimately affects attendance and test scores (Klem & Connell, 2006). Tutors who have firsthand knowledge of the curriculum and an ongoing established relationship with a student may more easily create and sustain a positive bond with the student than a contracted provider who has only limited contact with the student. The classroom teachers in this study were very familiar with the students and the curriculum. These teachers worked with the curriculum each day and had daily contact with the students.

Poor student attendance is consistently cited as a factor that contributes to poor outcomes of tutoring programs (Hock et al., 2001), and district teachers may be better positioned to influence attendance rates. With 100 % attendance as their goal, tutors in the current study carefully monitored students and requested homeroom teachers to remind students on tutoring days, thereby increasing the likelihood of attendance. Therefore, students received the necessary academic instruction, and tutors and teachers conveyed the message that they cared about their academic outcomes. Outside tutors do not have the same intimate connection with the students or school, thereby potentially limiting their influence on consistent attendance.

Klem and Connell (2004) found that as students believe a teacher is personally invested in their growth, engagement and attendance increase. The district teacher is viewed by the student as a school representative (Black, 2003). When a classroom teacher volunteers to serve as a tutor after work hours, students perceive this as a personal investment in the students' well-being, and it signifies the teacher's belief in the students' abilities. In this study, when the students were solicited for participation in the program, they were informed that they had been selected specifically because they were the group most likely to pass the test with additional tutoring.

Tutors in the current study demonstrated a strong belief in each student's capabilities, which potentially increased the student's confidence and motivation. This positive interaction in the tutoring sessions could transfer to the classroom environment. These factors may foster greater engagement in both the tutoring session and in school. This increase in confidence and engagement is ultimately connected to increased test scores (Klem & Connell, 2004).

This study's results suggest that students within this particular school benefitted from the tutoring program and point to some interesting questions for further research. Language arts and math tutoring by district teachers for students who enter tutoring at the near-passing range may produce positive student outcomes. Given the sample for this study, however, further investigation is warranted to determine if other borderline students would benefit equally from a similar program.

Limitations

Additional research should explore the specific components of district tutoring that affect student performance on standardized tests. For example, did students improve merely because they received additional academic time, or was it the content of the program that made a difference? This study did not evaluate specific components of the curriculum used during the tutoring session. The curriculum for the tutoring sessions was a standardized test preparation program. It would be interesting to replicate the study using different types of curricula to see if there are significant differences achieved based on the type of curriculum used for the tutoring sessions.

Identifying effective components of after-school programs is difficult for many reasons. Most samples used for program evaluations include selection bias, and Slavin (2008) cautioned that selection bias is

particularly problematic in studies of after-school programs. Because most after-school programs are voluntary, students who enroll may be more motivated than those who do not; thus, equating treatment and control groups is problematic (Fashola, 1998). Fashola's review of past evaluation studies found only two programs that used random assignment. Ritter and colleagues' (2009) meta-analysis included 21 studies that had used randomized field trials; however, they reported very small samples.

In the current study, due to the small sample of volunteers, the researchers were precluded from randomly assigning participants to the treatment group. For this reason, the study included non-volunteers who met the same eligibility criteria as the control group. While the researchers used ANCOVA to account for initial differences on standardized test performance, there are possibly inherent differences between our groups.

Other potential barriers to validity include sporadic attendance and fidelity of implementation, signifying the need for proper staff training (Fashola, 1998). In this particular study, the attendance rate was 90%. Furthermore, tutors used a test prep curriculum and followed the protocols they were expected to use every day in their regular classrooms, making the tutoring program an extension of the classroom program. Teachers used the test prep booklets, and the students completed one chapter per week; teachers were instructed to reinforce the test-taking strategies as defined by Standard Solutions (2006). Therefore, the researchers assumed that the teacher followed the program as specified.

This study is limited by the small sample size and needs to be replicated with a larger sample. The study is also limited by the student population selected for this study, as participants were not randomly selected but were invited into the program if they had previously scored at the near-passing range on standardized tests. Future research should include students scoring in the bottom and top range on previous standardized tests.

Conclusion

Research on middle school success illustrates the importance of relationships among students, teachers, administrators, and community (Feldman & Ouimette, 2004; Jackson, 2009) and establishes high expectations for all students (Cassellius, 2006). In the current study, the use of in-district tutors affected

student performance, perhaps due to the nature of the teacher-student relationship, tutors frequently communicated with students' teachers about their performance and instructional needs. The nature of this relationship, the acknowledgement of student effort, and establishment of high expectations (e.g., conveying the message from the beginning that they were selected due to the likelihood of success) may have contributed to student success. Whether the curriculum affected student success is unknown, but having caring, committed educators appears to influence student engagement and success.

References

- Baker, J., Rieg, S., & Ckendaniel, T. (2006). An investigation of an after school math tutoring program: University tutors + elementary students = a successful partnership. *Education*, *127*(2), 287–293.
- Black, S. (2003). Keeping kids from dropping out. *Educational Digest*, *68*(5), 37–42.
- Borja, R. (2007). Companies want changes in NCLB tutoring policies. *Education Week*, *26*(20), 10.
- Burns, M. K., Senesac, B. V., & Symington, T. (2004). The effectiveness of the HOSTS program in improving the reading achievement of children at-risk for reading failure. *Reading Research and Instruction*, *43*(2), 87–103.
- Cassellius, B. (2006). Using relationships, responsibility, and respect to get from “good to great” in Memphis middle schools. *Middle School Journal*, *37*(5), 4–15.
- Cooper, H., Charlton, K., Valentine, J. C., & Muhlenbruck, L. (2000). Making the most of summer school: A meta-analytic and narrative review. *Monographs of the Society for Research in Child Development*, *65*(1), 1–118.
- Davis, M. (2006). Report: Schools could improve on NCLB tutoring, choice. *Education Week*, *25*, 31.
- Fashola, O. S. (1998). *Review of extended-day and after-school programs and their effectiveness*. Baltimore, MD: Center for Research on Education of Students Placed At Risk CRESPAR Report No. 24). Retrieved from: www.csos.jhu.edu/crespar/techReports/Report24.pdf
- Feldman, J., & Ouimette, M. (2004, April). *Examining the Turning Points comprehensive middle school reform model: The role of local context and innovation*. Paper presented at the Annual American Educational Research Association Conference, San Diego, CA.

- Gordon, E. E. (2009). Five ways to improve tutoring programs: Evidence on tutoring points to practices that are found in the most successful tutoring programs. *Phi Delta Kappan*, 90, 440–445.
- Gordon, E., Morgan, R., Ponticell, J., & O'Malley, C. (2004). Tutoring solutions for No Child Left Behind: Research, practice, and policy implications. *NASSP Bulletin*, 88(638), 59.
- Hock, M. F., Pulvers, K., Deshler, D., & Schumaker, J. (2001). The effects of an after-school tutoring program on the academic performance of at-risk students and students with LD. *Remedial & Special Education*, 22, 172–186.
- Jackson, A. (2009). New middle schools for new futures. *Middle School Journal*, 40(5), 6–10.
- Klem, A. M., & Connell, J. P. (2004). Relationships matter: linking teacher support to student engagement and achievement. *Journal of School Health*, 74, 262–273.
- Lauer, P. A., Akiba, M., Wilkerson, S. B., Apthorp, H. S., Snow, D., & Martin-Glenn, M. L. (2006). Out-of-school time programs: A meta-analysis of effects of at-risk students. *Review of Educational Research*, 76, 275–313.
- Lewis, A. (2006). Washington scene. *Education Digest*, 72(2), 70–73.
- McComb, E. M., & Scott-Little, C. (2003). *After-school programs: Evaluations and outcomes*. Greensboro, NC: The Regional Educational Laboratory at SERVE.
- Morris, D. (2005). A Title I reading success story. *Reading Research and Instruction*, 45(1), 1–17.
- O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K–12 curriculum intervention research. *Review of Educational Research*, 78(1), 33–84.
- Pascopella, A. (2004). Signs of improvement with SES. *District Administration*, 40(6), 21–21.
- Preparing for the New Jersey Grade Eight Proficiency Assessment* (GEPA). (2001). Circle Pines, MN: American Guidance Service.
- Ritter, G. W., Barnett, J. H., Denny, G. S., & Albin, G. R. (2009). The effectiveness of volunteer tutoring programs for elementary and middle school students: A meta-analysis. *Review of Educational Research*, 79(1), 3–38.
- Skinner, E. A., & Belmont, M. J. (1993). Motivation in the classroom: Reciprocal effects of teacher behavior and student engagement across the school year. *Journal of Educational Psychology*, 85, 571–581.
- Slavin, R. (1999). How can funding equity ensure enhanced achievement? *Journal of Educational Finance*, 24, 519–528.
- Slavin, R. (2008). Perspectives on evidence-based research in education: What works? *Educational Researcher*, 37, 5–14.
- Standard Solutions. (2006). Available at <http://www.standardssolution.com/index.cfm?CFID=270198&CFTOKEN=66217646&fuseaction=edsolConsolHome.instruction>
- Triplett, C. F. (2004). Looking for a struggle: Exploring the emotions of a middle school reader. *Journal of Adolescent and Adult Literacy*, 48, 214–222.
- Zuelke, D. C., & Nelson, G. J. (2001). The effect of a community agency's after-school tutoring program on reading and math GPA for at-risk tutored students. *Education*, 121, 799–809.

Author's Note

Correspondence should be sent to Dr. Terri Rothman, Monmouth University, School of Education, 400 Cedar Avenue, West Long Branch, NJ, 07764, USA (732) 571-7507, trothman@monmouth.edu