New Evidence that Tutoring with Community Volunteers Can Help Middle School Students Improve their Academic Achievement

Anna Allen and Nancy Feyl Chavkin

Abstract

This study evaluates the impact of minimally trained tutors on the pass/fail ratio of middle school students in core subjects such as English, reading, math, or science. Because a traditional experimental design comparing students who receive a treatment with those who do not was not financially feasible, the study employed an alternative that has proven very successful when traditional control groups are not available. The within-program control group alternative permits a comparison between those who receive more of a treatment and those who receive less. For this evaluation, students were divided into two groups based on the number of hours of tutoring they received. Tutored students were more likely to increase their grade to passing than could be expected by chance. Students in the group that received more tutoring were more likely to pass than those who received less tutoring. The within-program control group design gives us confidence that tutoring of middle school students by minimally trained community volunteers can be effective.

Key Words: tutors, community volunteers, middle school, student academic achievement

Introduction

Dropping out of school continues to be a major national problem. Recent statistics from the National Dropout Prevention Center (2004) confirm the grim reality that dropouts are costing the country billions during their lifetimes in lost earnings and unrealized tax revenue. Dropouts earn less and are unemployed at a much higher rate than high school graduates. Harlow (2003) reports that 75% of state prison inmates are high school dropouts. Students with low academic achievement are twice as likely to become parents in high school and not complete their degree (Alliance for Excellent Education, 2003). Staresina (2004) suggests that because there are so many long-term consequences for students who drop out of school, it is imperative that we implement strategies such as tutoring programs to improve students' opportunities to complete high school.

This research reports on the results of a tutoring program that uses minimally trained local volunteers to tutor middle school students who are at risk for dropping out of school. It provides an opportunity to examine a low-cost, highly effective way of using community volunteers to improve the academic achievement of middle school students.

Review of Literature

Dynarski and Gleason (1999) identified several promising program models and strategies that helped reduce the dropout rate. The most successful of these programs were usually conducted for students in middle grades or earlier and were focused on specific students with targeted goals.

There is little debate that tutoring for middle school students who are at risk for dropping out of school seems to be a good idea and is in accord with the findings about best practices for keeping students in school. McElvain and Caplan (2001) report that schools and communities see great hope in tutoring programs for middle school students; however, there has been little evidence to date that tutoring programs can work beyond the early grades and even less evidence that tutoring is effective using minimally trained community volunteers during the regular school day. Most of the research on successful tutoring programs either focuses on early reading tutoring (Slavin, Karweit, & Madden, 1989; Wasik & Slavin, 1993) or reports on after-school programs that combine mentoring and tutoring (Cardenas, 1989; Huang, 2001; Leto, 1995). Programs with one-to-one tutoring using certified teachers appear to have larger effects than those using paraprofessionals or volunteers (Wasik, 1997), and thus the issue of cost is raised.

Volunteer tutors provide individual attention that most teachers cannot provide, especially as class size expands in response to tighter school budgets. Wasik and Slavin (1993) have found that one-to-one tutoring is the most effective individual remedial strategy if implemented well. Tutoring for reading, especially for beginning readers, has been studied far more than tutoring for other subjects. Most studies of early readers conclude that tutors with more training and expertise have a greater impact (Cohen, Kulik, & Kulik, 1982; Shanahan, 1998; Wasik 1998; Wasik & Slavin, 1993). These programs, however, are generally expensive and the samples tend to be small.

Two recent studies that looked at tutoring programs for "at-risk" beginning readers found significant improvement using college students with minimal training (Fitzgerald, 2001) and community volunteers (Baker, Gersten, & Keating, 2000). The Fitzgerald study using college students also found a significant correlation between hours of contact and successful outcomes. Middle school students are also represented in the tutoring literature, but in much smaller numbers. One school district in Florida, however, discontinued its summer school for middle school students who failed subjects during the academic year and replaced it with a successful after-school tutoring program (Lacey, LeBlan, & Collins, 2000). The evaluation focused on implementation rather than outcomes.

Reviews of effective reading tutoring programs identified a number of features that are critical to a successful program (Wasik, 1998; Moss, Swatz, Obeidallah, Stewart, & Greene, 2001):

- intensity of tutoring—frequency, session length, and individualized;
- structured sessions;
- close coordination with teacher and classroom;
- extensive tutor training—before and during course of tutoring; and
- careful monitoring of the effectiveness of tutoring services.

Description of Middle School Tutoring Program

Building on the recommendations from the literature and because meta-analyses of the dropout literature (Jimerson, 2001; Jimerson, Anderson, & Whipple 2002) strongly suggest that grade retention is one of the most powerful predictors of dropout status, this tutoring program was established for students who were experiencing serious difficulty passing a course and thus at most risk. The tutoring program was part of a larger dropout prevention initiative of Community In Schools in central Texas where an AmeriCorps program is responsible for tutoring and mentoring middle school students. This study focuses on the tutoring component of the AmeriCorps program only.

The goal of the tutoring program was to assist students to pass core subjects (math, reading or English Language Arts, science) for which they had received a failing grade the previous year. Each participating school received one to four volunteer tutors who could elect to work either full- or part-time. Community In Schools' on-site program managers supervised the tutors and were available for consultations on an as-needed basis. Program managers were primarily social workers who helped the tutors understand the academic, social, and emotional needs of students.

Teachers referred students to the tutoring program if they had failed a six-week grading period in the following subjects: math, English, reading, science, or social studies. Each tutor and student developed an individual plan for when and where tutoring would take place. Some students were tutored exclusively after school, others during class or in the library during an elective class such as band. Some focused on homework, others on practicing skills such as multiplication or reading comprehension. Tutors were expected to tutor for at least one hour a week, but not all of the students received the same number of hours of tutoring.

Methodology

Participants in the program included tutors and students from one large urban and two smaller rural school districts. The majority of the participating campuses were inner-city schools, and all schools served a high proportion of low-income families.

Tutors

The 31 AmeriCorps tutors were a diverse group, though 84% were female. Fifty-eight percent were of European American heritage, 19% were African American, 23% were Hispanic, and 4% were of Asian American heritage. College graduates comprised only 13%, 52% had some college, 32% were high school graduates, and the remaining 6% earned a GED. The attrition rate for the volunteer tutoring program was 27%, considered to be within the normal range for national AmeriCorps programs.

Most of the 31 AmeriCorps tutors began their service in September and October during the 2001-2002 school year and received five days of preservice training. The members who signed up later did not receive the same systematic preparation and training. Five days of preservice training included one day devoted to learning how to tutor math and reading. Training to tutor was necessarily general, because tutors would be addressing many different subjects and three grades. Other days of training covered subjects such as how to work

in a school, working with early adolescents, and mentoring. Training continued throughout the year with approximately four more hours of training in tutoring, especially math. Program managers provided on-the-job training in mentorship as well as tutoring, and they trained the tutors to encourage students to talk about any problems that seemed to keep them from focusing on their school work. Volunteer tutors met together regularly to share successes and learn more about how to tutor.

Students

The student sample is comprised of 256 middle school students (6th through 8th grades). Of the tutored students, 61% were male; 79% of the tutored population was eligible for the free/reduced lunch program. Hispanic students comprised 61% of those tutored, 22% were African American, 13% were of European American heritage, 3% were other, and 1% did not list ethnicity. Eighteen percent of the students were from homes in which English is not the primary language. Many of the tutored students were reported by their teachers to be working below grade level and/or have limited English skills. Approximately 88% of the students met the most stringent criteria for inclusion in the tutoring program—a failing six-week report card grade or failure in a core subject the previous year. The rest of the sample, 12%, was at risk of failing a course because of borderline grades (70-73%), according to their teachers.

Procedures

In the classic experimental-control comparison, the control group does not receive any of the treatment. In this way the effectiveness of a treatment can be measured by comparing those who received it with a similar group who did not. The expense and difficulty associated with securing traditional control groups in three different school districts prohibited a traditional experimental design for this evaluation. The wide range of treatment received by the students in this evaluation suggested an alternative, however. Control groups are frequently difficult to work out in public school settings, so researchers in the field of education have devised an alternative design when a traditional control group is not feasible (Invernizzi, Rosemary, Juel, & Richards, 1997; Fitzgerald, 2001). In this alternative method, a within-program control group is formed when a group of children is compared with a group of similar children who received the same treatment, but less of it. The alternative design essentially tests the effectiveness of a treatment by comparing those who received more of it to those who received less of it. "This alternative within-program control group is potentially a more stringent test of a treatment than the classic form

of comparing something to nothing. If statistically significant outcomes arise, they clearly can be directly attributable to what is happening in the program in conjunction with the amount of treatment that is provided" (Fitzgerald, 2001).

Volunteers used a computerized data collection system to record the type of service and hours of contact they had with their students. Time that the volunteer spent that was not direct contact with the student was not recorded for this evaluation. Thus time spent recording information or checking on grades or looking for a student was omitted. The intent was to measure only the time that the volunteer actually spent with their students.

Progress was measured by comparing the previous year's failing grade in a core subject (English, reading, math) or a failing six-week report card grade with the final grade the student received at the end of the current school year in the targeted subject.

The students were divided into two groups based on the distribution of the number of hours of tutoring they received. Thus, half of the students received 13.25 hours of tutoring and the other half received 13.50 to 61 hours of tutoring. The two groups were similar in every other respect, but one group received more tutoring than the other. The primary reason for the difference between the two groups is that the group receiving less tutoring started tutoring later in the year because they did not fail in the first or second six-week grading periods.

Results

Figure 1 shows that the average reported six-week grade before tutoring was 60; the average grade at the end of the year was 73. The final grade at the end of the year includes the previous lower grade. In many instances, the tutoring did not begin until after the student had failed more than one six-week grading period. Ethnicity and receiving free/reduced lunch did not affect the results, but gender did. As in other studies, females tended to be more likely to pass than males ($x = 2.045$, $df = 1$, $p = .004$). The paired t-test was used in order to compare the mean of the initial six-week report card grade with the mean of the final grade in the subject for which the student received tutoring. A paired t-test showed that the tutored students made significant improvement from their initial six-week grade to the grade at the end of the year ($t = 19.254$, $df = 216$, $p < 0.00025$, two-tailed).

Researchers also look at the results to see if there were any differences in pre- and post-grades by subject or by grade level. There were no significant differences.

Figure 1. Comparison of Average Pre-Grade with Final Grade

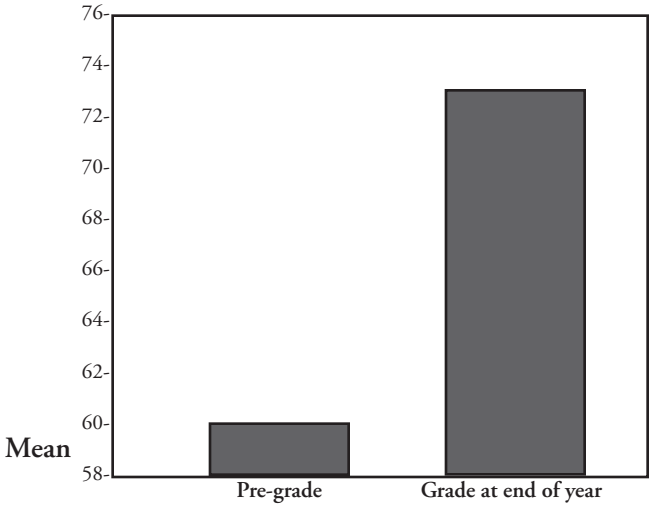
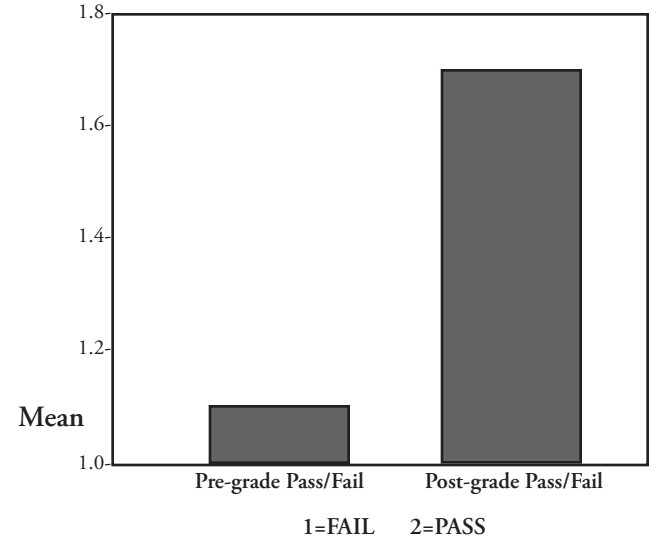


Figure 2 illustrates the rate of pass/fail where one equals a failed grade and two equals a passing grade. Because the bars represent the average rates of passing or failing, the range for the mean pass/fail score is approximately 1.1 to 1.7. Figure 2 includes students whose pre-grade was passing and whose final grade was failing.

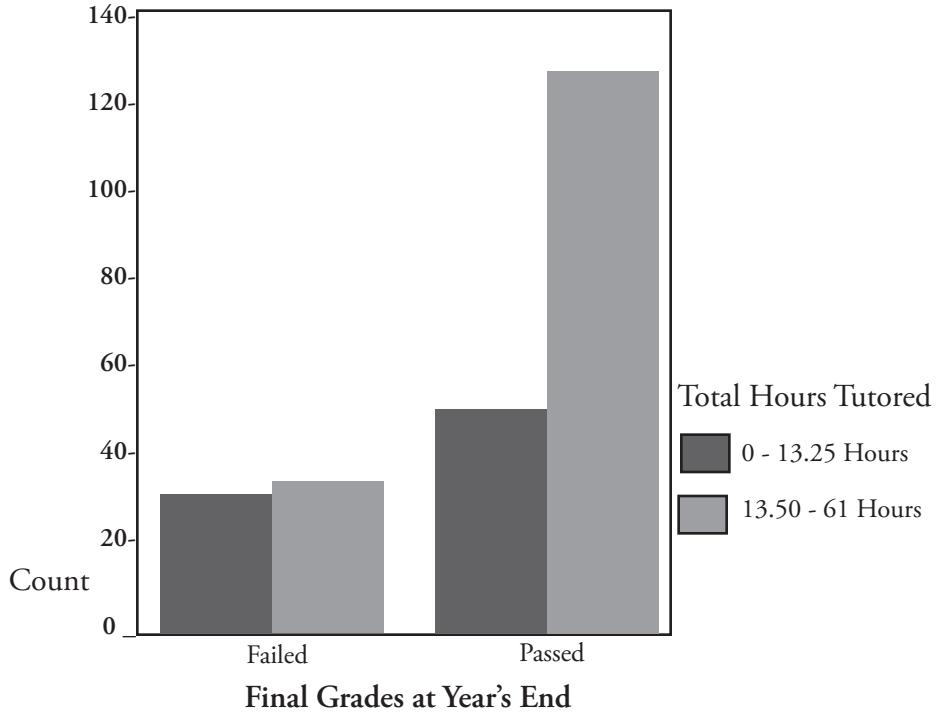
Figure 2. Comparison of Average Rate of Pass/Fail Before and After Tutoring



Again, researchers examined the results to see if there were any differences in the average pre- and post- pass/fail scores by subject or grade level. There were no significant differences.

Figure 3 shows the results when the tutored students were divided into two groups based on the amount of tutoring they received. Approximately half of the students (120) in the tutoring program sample received between 1 and 13.50 hours of tutoring. The other half (126) received between 14 and 61 hours of tutoring. To test for tutoring effects on the rate of pass/fail for end of the year grades in either math, language arts, reading, science, or history, students who received less than 14 hours were compared to the second group who received 14 or more hours of tutoring. In this comparison (Figure 3), students who received more tutoring were compared to those who received less.

Figure 3. Tendency Toward Pass/Fail by Students Who Received More Tutoring Compared with Students Who Received Less Tutoring



The two groups are similar across different grades, schools, and districts. Gender, age, poverty, and ethnicity were tested for contributions to differences between the two groups, and the only significant difference was for free/reduced lunch ($\chi^2 = 6.006$, $df = 1$, $p = 0.014$).

The chi-square test is used to compare the distribution of frequency data collected in a study with the distribution that would be expected to occur by chance. This statistic suggests the impact of the number of hours on the final grade. In this study, we compared the frequency of failure rates of students who received more tutoring with the failure rates of students who received less tutoring. As Table 1 indicates, students who received less tutoring were significantly more likely to fail; those who received more tutoring were significantly more likely to pass ($\chi^2 = 9.71$, $df = 1$, $p = 0.002$).

Table 1. Tendency Toward Pass/Fail by Hours of Tutoring Cross Tabulation

Tendency Toward Pass/Fail Rate in Tutored Subject		Total Hours Tutored		Totals
		0 – 13.25 Hrs	14 – 61 Hrs	
Failed	**Count**	29	31	60
	%	39.2%	19.9%	26.1%
Passed	**Count**	45	125	170
	%	60.8%	80.1%	73.9%
Total	**Count**	74	156	230
	%	100.0%	100.0%	100.0%

Another test of correlation was used to determine the size effect or strength of the correlation between the likelihood of passing or failing and hours tutored. This statistic helps answer the question of how much the tutoring contributed toward the increased number of students who passed. Pearson's correlation was significant, with a respectable 21% effects size ($r = .212$, $n = 255$, $p = 0.001$, two-tailed).

Discussion

The results of the tutoring program are very promising. As both Figure 1 and Figure 2 demonstrate, there is a significant increase in the number of students who achieved both a higher grade and a passing grade at the end of the year compared to the grade they had before tutoring began. There were no significant differences in subject areas or grade levels.

In Figure 3, there was a higher tendency toward passing for students who received more tutoring. After testing for gender, age, poverty, and ethnicity for any significant differences between the two groups, a significant difference did emerge for free/reduced lunch. This difference is primarily attributable to the rural schools, who have a higher proportion of students who do not receive

free/reduced lunch than the more urban schools and whose volunteers delivered an impressive number of hours to their tutored students. Other than the rural schools, however, students who received less tutoring are from the same schools as those who received more tutoring. In many instances, the students who had less tutoring had the same tutors as those who received more tutoring. In sum, the within-program control group appears to be a good, though not perfect, comparison group.

The within-program control group alternative evaluation shows that the more tutoring students receive, the more likely they will pass the course for which they are receiving tutoring. Other data supports this conclusion as well. For example, those who enter the program after the first six weeks are significantly ($p < .005$) more likely to pass than those who enter the program later in the year. The comparison is weakened somewhat by the very slight overrepresentation of students ineligible for free/reduced lunches in the group receiving more tutoring; however, as noted earlier, the number of students who do not receive free/reduced lunch is somewhat mushy, because if nothing was recorded for a student, he or she was recorded as not eligible for free/reduced lunch. Many students were also receiving other types of public assistance, but did not apply for free/reduced lunch. This is especially true for intermediate students.

More studies are needed to examine the relationship between tutoring and academic achievement. Although this study has produced some interesting results supporting the positive effects of tutoring, it is important to recognize the limitations of this study and to make suggestions for future studies. Data were not gathered that differentiated tutoring approaches. Future studies might discriminate between helping with homework and practicing skills such as reading or multiplication tables. There might be a difference among tutoring approaches. Another factor that might be examined is the difference between tutoring small groups of two to three students and one-to-one tutoring.

Future studies may also want to examine the interaction between tutor and teacher. The information gathered in this study was not systematic enough to draw any conclusions, but there was strong indication that there were differences in relationships between tutors and teachers. Some teachers worked closely with the tutors, and others interacted infrequently with the tutors.

Conclusion

Although more studies need to be conducted, the within-program comparison group strongly suggests that tutoring by minimally trained tutor/mentors increased the likelihood that a student would pass a core subject. The tutoring

program garnered many very positive comments from teachers who praised the program for "turning students around."

The results of this study have important implications for community after-school programs and volunteer tutoring of middle school students. At least 14 hours of tutoring were required to help a student move from a failing grade to a passing grade. Minimally trained adult volunteers without a college education can provide enough academic support to increase the rate of passing for middle school students.

This study lends support to the significant role that community volunteers can play in increasing the academic achievement of middle school students. Funding for programs that link community volunteers with students at risk of dropping out of school can make a difference in helping students pass their courses and stay in school.

References

Alliance for Excellent Education. (2003). *Fact Sheet: The impact of education on health and well-being*. Washington, DC: Author.

Baker, S., Gersten, R., & Keating, T. (2000). When less may be more: A 2-year longitudinal evaluation of a volunteer tutoring program requiring minimal training. *Reading Research Quarterly, 35*, 494-519.

Cardenas, J. (1989). *Report on the Valued Youth Partnership of San Antonio*. San Antonio, TX: Intercultural Development Research Association, Center for the Prevention and Recovery of Dropouts.

Cohen, P. A., Kulik., J. A., & Kulik, C. C. (1982). Educational outcomes of tutoring: A meta-analysis of findings. *American Educational Research Journal, 19*, 237-248.

Dynarski, M., & Gleason, P. (1999). *How can we help? Lessons from federal dropout prevention programs*. Princeton, NJ: Mathematica Policy Research.

Fitzgerald, J. (2001). Can minimally trained college student volunteers help young at-risk children to read better? *Reading Research Quarterly, 36*(1), 28-47.

Harlow, C. W. (2003). *Education and correctional populations. Bureau of Justice special report*. Washington, DC: U.S. Department of Justice.

Jimerson, S. R. (2001). Meta-analysis of grade retention research: Implications for practice in the 21st century. *School Psychology Review, 30*(3), 420-437.

Jimerson, S. R., Anderson, B., & Whipple, A. (2002). Winning the battle and losing the war: Examining the relation between grade retention and dropping out of high school. *Psychology in the Schools, 39*(4), 441-457.

Huang, D. (2001). An after-school evaluation system for middle and high school programs. *NASSP Bulletin, 85.626*, 45-61.

Invernizzi, M., Rosemary, C., Juel, C., & Richards, H. C. (1997). At-risk readers and community volunteers: A 3-year perspective. *Scientific Studies of Reading, 1*, 277-300.

Lacey, C., LeBlanc, P., & Collins, M. (2000). Program evaluation: Lessons to be learned from a district-wide middle school tutorial program. *Planning and Changing, 31*, 166-181.

Leto, D. J. (1995). Focus on research: Creating community with an after-school tutoring program. *Language Arts, 72*(2), 128-136.

McElvain, C. K., & Caplan, J. C. (2001). Creating effective after-school programs for middle and high school students. *NASSP Bulletin, 85.626*, 35-44.

Moss, M., Swatz, J., Obeidallah, D., Stewart, G., & Greene, D. (2001). *AmeriCorps tutoring outcomes study*. Retrieved May 20, 2004 from http://www.americorps.org/research/pdf/tutoring_0201.pdf

National Assessment of Educational Progress (NAEP). (1984/1992). *The reading report card*. Washington, DC: Author.

National Dropout Prevention Center. (2004). *Quick facts*. Retrieved June 5, 2004 from http://www.dropoutprevention.org/stats/quick_facts/econ_Impact.htm

Shanahan, T. (1998). On the effectiveness and limitations of tutoring in reading. In P. D. Pearson & A. Iran-Nejad (Eds.), *Review of research in education* (pp. 217-234). Washington, DC: American Educational Research Association.

Silverman, D. (1993). *Interpreting qualitative data: Methods for analyzing talk, text, and interaction.* London: Sage.

Slavin, R. E., Karweit, N. L., & Madden, N. A. (Eds.). (1989). *Effective programs for students at-risk.* Boston: Allyn and Bacon.

Staresina, L. N. (2004). Educational Issues A-Z: Dropouts. *Education Week on the Web*. Retrieved July 10, 2004, from http://www.edweek.org/context/topics/Issuespage.cfm?id=113

Wasik, B. A. (1997). Volunteer tutoring programs in reading: Do we know what works? *Phi Delta Kappan, 79*, 282-7.

Wasik, B. A. (1998). Volunteer tutoring programs in reading: A review. *Reading Research Quarterly, 33*, 266-292.

Wasik, B. A., & Slavin, R. E. (1993). Preventing early reading failure with one-to-one tutoring: A review of five programs. *Reading Research Quarterly, 28*, 178-200.

Anna Allen is an independent evaluator in the Austin, Texas area, and she has extensive experience designing and conducting evaluations of school and community programs.

Nancy Feyl Chavkin is the director of the Richter Institute of Social Work Research and professor at Texas State University-San Marcos, and she has published in the areas of family and community involvement in education, school social work, child welfare, and research utilization. Correspondence concerning this article should be addressed to Prof. Nancy F. Chavkin, Center for Children & Families, Texas State University-San Marcos, 601 University Dr., San Marcos, TX, 78666 or to nancychavkin@txstate.edu.