Applying Matched Sampling to Evaluate a University Tutoring Program for First-Year Students

Mark E. Walvoord[1] and Jacob D. Pleitz[2]
University of Oklahoma

Author's Note

[1]*Now at the University of Central Oklahoma, 100 North University Drive, Box 212, Edmond, OK 73034.*
[2]*Now at Virginia Polytechnic Institute and State University*

Abstract

Our study used a case-control matching design to assess the influence of a voluntary tutoring program in improving first-year students' Grade Point Averages (GPA). To evaluate program effectiveness, we applied case-control matching to obtain 215 pairs of students with or without participation in tutoring, but matched on high school GPA and standardized test scores. Next, we examined differences in academic performance between the two groups. A matched pairs t-test showed that students who attended the tutoring sessions demonstrated significantly higher GPAs during their first year. We close by discussing implications and suggestions for future research.

Introduction

Cross-Level Peer Tutoring

Peer tutoring has long been used in education as an effective learning tool. With settings ranging from elementary school (Ladd & Kochenderfer, 1996) to higher education (Colvin, 2007), the practice of students helping to teach other students is often praised in the literature as a method for increasing student motivation (Miller & MacGilchrest, 1996), improving student learning (Millis & Cottell,

1998; Entwistle, 1997), and fostering greater student academic responsibility (Goodlad, 1998). Given the effectiveness of such practices, many schools are adding peer tutoring as a method of improving student performance.

Peer tutoring differs from traditional lecture methods by employing peers, rather than professional teachers, as the primary means of instruction. Although it may seem counterintuitive that a less-experienced expert may be a more effective teacher, research results present many reasons why this may be the case. We'll start by defining what exactly a peer is.

According to Falchikov (2001, p. 1), a peer is someone of "the same social standing," and peer-level tutoring involves "helping each other to learn and learning themselves by teaching" (Topping, 1996; Falchikov, 2001, p. 3). Damon and Phelps provide a more specific definition of this process: "Peer tutoring is an approach in which one child instructs another child in material on which the first is an expert and the second is a novice" (1989, p. 11). Two important points may be derived from this definition, specifically highlighting why peer-level tutoring may be more a more effective method of learning than traditional instructor-based learning.

First, because peer tutoring involves teachers and learners of the same social standing, there may exist a unique social and intellectual reciprocity between the two. Since the teacher and learner are on the same peer level, they may be more likely to share a bond not experienced by the traditional student-professor dynamic. Given that both roles share a similar social status, they may be more likely to identify with the troubles and difficulties of the material in a unique perspective not shared by a more proficient expert. Drawing along the lines of Piaget's learning theory, this shared bond may help the student and learner work together to develop more similar schemata, or building blocks, towards understanding concepts (Piaget & Cook, 1952).

Second, because peer-level tutoring allows students to work together in the teacher-learner dynamic, some people have proposed that a more appropriate term, mutual instruction, be used to describe the process (Swengel, 1991). This term helps emphasize the second potential advantage of peer level tutoring—that by teaching the

material to their classmates, students are able to learn and master the material from a different perspective. By both teaching and learning from one another, students are allowed to engage in a form of cooperative learning (Hermann, 2013; Machemer & Crawford, 2007), a form of learning connected with many positive outcomes, including retaining information longer (Lujan & DiCarlo, 2006) and increased academic confidence (Cherney, 2008).

There are two major characteristics that should differentiate the peer tutor from the "client" student. Tutors must have high-quality training (Sheets, 1994) and have clearly defined roles as tutors instead of course teaching assistants (Carsrud, 1979) to be most effective. Proper training gives the tutor confidence in their abilities, skills in communication, and a background in learning theory that allows them to be most effective in working with peers. Unlike teaching assistants, tutors don't have control over students' grades or the construction of course materials, so clients view them as a more neutral resource for discussing course challenges.

Barbara Millis, in Cooperative Learning in Higher Education (2010), summarizes the usefulness of cooperative learning, even apart from any peer tutoring constructs, in effectively encouraging deep learning, critical thinking, and academic skill-building. Properly structured group learning opportunities serve the purpose of increasing topical understanding and academic skills, building communities to connect students to the institution, and increasing a sense of belonging leading to increased retention (Kuh, Kinzie, Buckley, Bridges, & Hayek, 2007).

Our Action Tutoring Program

This paper investigates a voluntary tutoring program at a large, public, mid-western university. Created in 2007, it is a campus-wide tutoring program designed to increase students' academic achievement, critical thinking skills, and positive study habits with the ultimate goal of helping to boost retention and graduation rates. Tutoring sessions are staffed by undergraduate Peer Learning Assistants (PLAs) who have previously completed the course that they are tutoring and attend College Reading and Learning Association certified training for quality assurance (College Reading & Learning Association, 2015). Alternatively, sessions may be staffed

by course graduate teaching assistants and/or by course instructors. Called U.C. Action, to promote the idea that students must be proactive in improving their educational experience, the program was originally designed to offer assistance in introductory courses with the highest rates of D, F, or W grades. Specific opportunities offered by the U.C. Action program include group tutoring sessions to encourage peer collaboration and learning communities, faculty lead activities to give the students greater exposure to their professors, and the use of learning enhancement tools such as SmartBoards, iPads, worksheets, and visual aids to integrate active learning.

Faculty participation is encouraged through $500 grant opportunities, increased interactions with students outside a large lecture setup, efficiency of using their office hours in this way, and the hope of increased student learning in their courses. Approximately 30 faculty, 20 GTAs, and 40 Peer Learning Assistants (PLAs) per semester work with U.C. Action to provide weekly drop-in and by-appointment assistance in over 75 courses. The majority of client visits occur at the voluntary, drop-in, group-learning sessions hosted by faculty, but additional visits occur at those hosted by PLAs and at small-group, by-appointment sessions that are available face-to-face or online. For the 2013–2014 academic year, 4,034 students visited sessions over 19,490 times, averaging 4.6 hours of assistance and 3.9 visits per student.

Recent research suggests that attending voluntary tutoring sessions may have a positive effect on student performance and retention, particularly when students perceive a benefit in doing so (Cavanaugh, 2011). In two 2010 studies, both cross-level, voluntary, one-to-one tutoring for at-risk students (Rheinheimer, Grace-Odeleye, Francois, & Kusorgbor, 2010) and drop-in, group tutoring (Cooper, 2010) correlated positively with students' retention rates and academic performance at U.S. public institutions. Additional positive effects were found by Arco, Fernandez, Espin, and Castro (2006) at a Spanish public university, suggesting that the effect may not be unique to domestic institutions. Cross-level, voluntary, group tutoring is also an effective means of increasing course grades when students used at least 9 hours per semester of tutoring according to Munley, Garvey, and McConnell (2010), and it is an effective means

to increase overall GPA and retention (Coladarci, Willett, & Allen, 2013).

Case-Control Matching

A well-known factor of proper experimental design is that the groups being compared be as close as possible on every factor other than the one of interest. Because of this, a component of true experimental design is the notion of randomization (the idea that every participant has an equal probability of being assigned to any group). Having equal probabilities of assignment will, in effect, make the groups equal on every factor other than the one being manipulated. As such, any difference in the dependent variable should be attributed to the independent variable.

The notion of case-control matching design is presented in Shadish, Cook, and Campbell (2001), who demonstrate the effectiveness of this quasi-experimental design across a number of fields, including sociology (Holland, 1986), epidemiology (Ahlbom & Norell, 1990), and psychology (Holland & Rubin, 1983). Built upon Rubin's Causal Model (Rubin, 1977), case-control design seeks to answer the question of what if an alternate event had occurred. According to Rubin, every one unit in an experiment has a potential outcome, depending on the condition to which the unit is assigned. The question then becomes, what if the person had not been assigned to this condition, but rather to another condition. All things being equal, should we observe a different outcome, then we can assume the cause of this outcome to be the different condition.

In light of this logic, sometimes called counterfactual conditioning, the current situation provides an ideal environment for case-control design. Specifically, students are presented with 1 of 2 potential realities: either attending or not attending one or more tutoring sessions. The question then becomes, what would have happened, had the attending student not attended, or what would have happened if an un-tutored student had participated. In traditional experimental design, this is created through the process of randomization (Fisher, 1935); however, because we cannot randomly assign participants to groups in the current case, we used a case-control setup.

Case-control designs create "equal groups" through matching on variables known to potentially influence the outcome. For example, college grades are often directly related to high school grades and standardized test scores (Pleitz, Terry, & Campbell, 2010). Given these pre-existing potential differences, along with the knowledge that such differences may cause a difference in the outcome that cannot be attributed directly to the variable under interest, case-control designs seek to match on these variables. They effectively wash out their influence and create equal groups. By controlling for all known confounding variables, case-control designs then allow the experimenter to answer the question of what would have happened, had an alternate reality occurred.

Purpose of the Current Study

As tutoring programs become more popular, researchers and educators are becoming more interested the effectiveness of such programs in improving students' academic performance. However, given that attendance at many of these tutoring sessions is strictly voluntary, true experimental design is typically not possible. This current article provides an example of how case-control designs can be used to assess the influence of these programs in a more effective manner. Using case-control designs, researchers and administrators can gain a greater insight into the efficacy of such programs, and hopefully better understand how to improve student performance in higher education.

Methods

Participants

Data were collected from approximately 3,939 students at a large mid-western university (University of Oklahoma IRB approval #3563). Although exact ages of participants were not available, it is likely that they were between the ages of 18 and 20, because we restricted our study to first-year students and our campus has a fairly traditional student population. Approximately 2,135 (54%) of the sample were female. Of the sample, approximately 14% of students attended at least one U.C. Action tutoring session. Table 1 presents an overview of the sample's descriptive statistics.

Table 1

Labels and Descriptive Statistics for Sample (N=3,939)

Variable	N	Mean	SD	Min	Max
ACT	3,939	25.97	3.90	14.00	36.00
High School GPA	3,939	3.65	0.29	2.22	4.00
Average Minutes per Visit	557	44.78	40.68	0.010	201.07
Minutes (Total)	557	82.10	42.92	0.01	240.00
First Semester GPA	3,939	2.9787	0.88	0.00	4.00

Variable	Label	Frequency	Percentage
Gender	Gender		
Male	1	1,804	45.80%
Female	2	2,315	54.20%
Action Tutoring			
Yes	1	557	14.14%
No	0	3,382	85.86%

In the current situation, the cases/controls are matched on high school GPA and standardized test scores. These variables have been shown to be highly associated with college performance in previous studies (Pleitz et al., 2010), and explain 20% of the variance in first year GPA within the current study. Table 2 presents the means and standard deviations of high school GPA, ACT score, and first year college GPA for the two groups. The control group represents the students who did not attend the U.C. Action tutoring sessions while the study group represents the students who attended one or more sessions.

Table 2

Descriptive Statistics for Case and Control Groups

Group	High School GPA	Standardized Test Score	First Year GPA
Control	3.634 (0.3045	26.040 (3.986)	2.925 (0.923)
Study	3.744 (0.233)	25.554 (3.318)	3.305 (0.548)

Table 2 reveals that the study group (those students who attended the tutoring sessions) entered into college with higher incoming high school GPAs but a slightly lower average standardized test score than the control group. Because of this, it may be that their higher first year GPAs are a reflection of greater academic credentials, rather than being due to attending the tutoring sessions. To account for this, we matched students between the two groups on these variables, to get a more representative picture of the influence of the U.C. Action program on first year GPA.

Results

After creating the criteria for matching (equal high school GPA's and standardized test scores), the data set contained 215 pairs of matched students. Given that the original study data set contained 557 first-year students that had visited tutoring, appropriately-matched pairs were found for 39% of those students. After matching, the mean high school GPA for the group was 3.74 and the mean standardized test (ACT) score was 26.30.

Having matched the samples on the relevant predictors, we then examined the mean differences in first year performance between the two groups. A paired sample t-test was conducted to compare the first year GPAs in the study and control conditions. Table 3 presents the results from the t-test where the mean variable represents the difference between the study group and the control group.

Table 3

Results from Case-Control Paired Samples t-test

N	Mean	Std Error	Df	t value	pr < t
215	0.2909	0.0664	213	4.37	< .0001

Results indicated that there was a significant difference between the scores for the attendees (M=3.29, Sd=0.57) and the control group (M=2.99, Sd=0.92), $t(214)=4.37$, $p < .0001$. Figure 1 presents the distribution of differences in college GPA between the two groups. The solid curve represents the data distribution using a standard normalized curve, while the dotted line represents the data distribution using a kernel density method. The box plot presents the

mean difference in college GPA between the two groups, along with a 95% confidence interval for this difference.

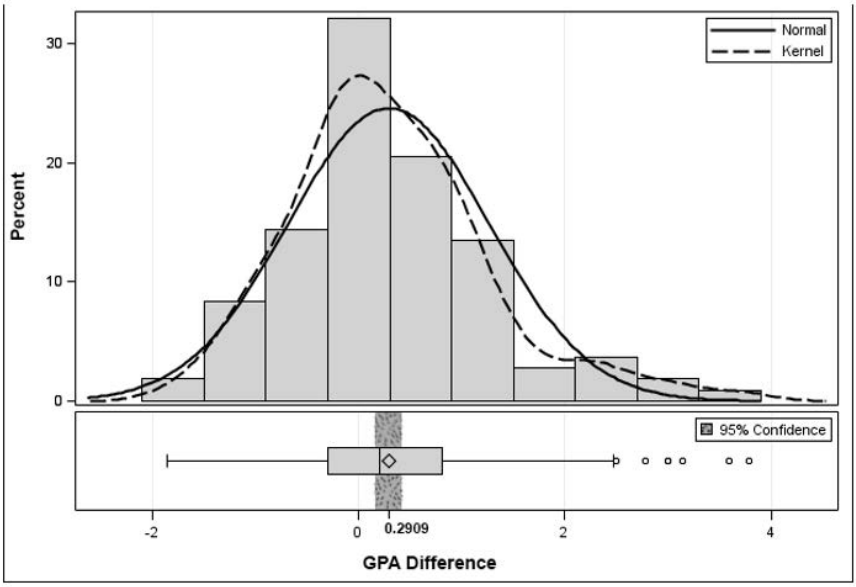

Figure 1. Distribution of differences in first year GPA between matched samples.

Discussion

The current study applied a case-control matched-pairs analysis to investigate the effectiveness of attending U.C. Action tutoring sessions on a student's first year academic performance. Students were matched on incoming high school GPAs and standardized test scores, and the differences in academic performance between the two groups were examined. Results indicated that the mean GPA for the group who attended at least one tutoring session was approximately 0.29 higher than the group who did not attend.

As we anticipated, those who attended at least one session had a higher average high school GPA than non-attendees, but we were surprised that their average standardized test scores were slightly lower. This implies that attendees are motivated to get better grades but may have some difficulty taking tests, and perhaps they attend tutoring hoping to increase test performance.

The results from the current study suggest that programs such as U.C. Action are having a positive effect on improving

students' first year GPAs. These findings continue to support the notion that implementation of peer-lead, and cross level tutoring may provide an effective method of improving student performance in higher education. Further, Given that many schools are feeling the increased pressures of budget cuts and faculty shortages, the current findings may be of particular importance within the current financial climate.

Peer tutoring can take many forms, and researchers continue to analyze the most effective methods for various educational settings. Topping (1996) summarized tutoring research in both K–12 education and higher education, showing evidence that peer tutoring is both an effective and cost-efficient means of increasing student performance, and in some cases, decreasing drop-out rates. Similarly, Falchikov (2001) gives several advantages to peer tutoring in higher education settings including decreased drop-out rates, increased study skills, and increased confidence. As the popularity of tutoring programs continues to grow, researchers have become more interested in investigating how these programs can increase retention, academic success, and graduation rates.

Along with demonstrating the effectiveness of cross-level, faculty-involved tutoring programs in improving student academic performance, the current results also present an effective way of implementing quasi-experimental design to assess cause and effect when true randomization is not a viable option. Specifically, by matching on factors known to influence the outcome of interest (student academic performance), the current article has demonstrated that the effects of random assignment may be achieved even in situations where participation in the experimental group variable is voluntary.

Shortcomings of the Current Study

Because the results presented above relied on post-hoc analysis of data observations rather than a full experimental setup, we must be careful with the interpretation of our results. It is important to point out that the students were neither randomly assigned to groups, nor randomly sampled from a general population. As such, it is not possible to directly state that participation in the tutoring program caused students to earn higher grades. Since our results were observed without any randomization or manipulation, any direct line

of causation is not possible due to potential confounds (Shadish et al., 2001).

Additionally, it is important to note that the above analysis is limited in scope due to its focus on the aggregate level of student body, rather than on the more microscopic details. That is, because students are lumped into one large sample, any nuanced behaviors are not detected. These include specific potential interactions between attendance and different majors, attendance at sessions with or without faculty present, and attendance and duration of visits. Although it appears that attending at least one session is helpful, it remains to be seen whether it is particularly more helpful for certain courses, majors, or types of students than others. These variables provide the greatest opportunity for future research.

Applications

While the present results indicate that participation in our voluntary tutoring sessions yields positive benefits for students, they also represent a pervasive problem long known to administrators and educators—that voluntary programs are only beneficial if students are willing to partake in them. Furthermore, as the above results indicate, students who are most likely to use the program are also those most likely to already have high grades to begin with. As such, it seems that those students who need academic assistance programs the most, may be the least likely to use them.

Understanding why lower achieving students are less likely to attend these effective sessions is an important first step to making such programs more attractive to target populations. Strategies we've used to make sessions more attractive and accessible to these target groups include having professors host the sessions, scheduling sessions at times and days to cater to students' needs and study habits, and offering sessions online to allow different learning preferences and flexibility of location.

Maximizing a student's academic potential is one of the most important goals for colleges and universities. This article has demonstrated how one method of instruction, a voluntary tutoring program, can be used to improve a student's first year GPA. Although there is a great deal more that needs to be studied within this area, including how to recruit low-achieving students to these services, by being able to demonstrate such programs' effectiveness,

researchers and administrators can provide students with at least one powerful reason for attendance.

References

Arco, J. L., Fernandez, F. D., Espin, A., & Castro, M. (2006). *A cross-age peer tutoring program to prevent academic failure and drop-out among first year university students.* Proceedings from the 36th ASEE/IEEE Frontiers in Education Conference, Session M2E, October 28–31, 2006, San Diego, CA.

Ahlbom, A., & Norell, S. (1990). *Introduction to modern epidemiology.* Chestnut Hill: MA, Epidemiology Resources.

Carsrud, A. L. (1979). Undergraduate tutors: Are they useful? *Teaching of Psychology, 6*(1), 46–49.

Cavanaugh, M. (2011). Students' experiences of active engagement through cooperative learning in lectures. *Active Learning in Higher Education, 12*(23), 23–33.

Cherney, I.D. (2008). The effects of active learning on students' memories for course content. *Active Learning in Higher Education, 9*, 152–171.

Coladarci, T., Willett, M. B., Allen, D. (2013). Tutor program participation: Effects on GPA and retention to the second year. *The Learning Assistance Review, 18*(2), 79–96.

College Reading & Learning Association. (2015). About ITTPC. Accessed 31 October 2015 at http://www.crla.net/index.php/certifications/ittpc-international-tutor-training-program.

Colvin, J. W. (2007). Peer tutoring and social dynamics in higher education. *Mentoring and Tutoring, 15*(2), 165–181.

Cooper, E. (2010). Tutoring center effectiveness: The effect of drop-In tutoring. *Journal of College Reading and Learning, 40*(2), 21–34.

Damon, W., & Phelps, E. (1989). Critical distinctions among three approaches. In N.M. Webb (Ed.), *Peer interaction, problem-solving and cognition: Multidisciplinary perspectives.* New York: Pergamon.

Entwistle, N. (1997). Reconstituting approaches to learning: A response to Webb. *Higher Education, 33*(2), 213–218.

Falchikov, N. (2001). *Learning together: Peer tutoring in higher education.* New York: Routledge Falmer.

Fisher, R. A. (1935). *The design of experiments.* Edinburgh: Oliver and Boyd.

Goodlad, S. (1998). Students as tutors and mentors. In S. Goodlad (Ed.), *Mentoring and tutoring by students* (pp. 1–17). London: Kogan Page.

Hermann, K. J. (2013). The impact of cooperative learning on student engagement: Results from an intervention. *Active Learning in Higher Education, 14*(3), 175–187.

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association, 81*(396), 945–960.

Holland, P. W., & Rubin, D. B. (1983). On Lord's paradox. In H. Wainer & S. Messick (Eds.), *Principals (sic) of modern psychological measurement* (pp. 3–25). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kuh, G. D., Kinzie, J., Buckley, J. A., Bridges, B. K., & Hayek, J. C. (2007). Piecing together the student success puzzle: Research, propositions, and recommendations. *ASHE Higher Education Report, 32*(5), 1–182.

Ladd, G. W., & Kochenderfer, B. J. (1996). Linkages between friendship and adjustment during early school transitions. In: W. M. Bukowski, A. F. Newcomb & W. W. Hartup (Eds.), *The*

company they keep: friendship in childhood and adolescence (pp. 322–345). Cambridge, UK: Cambridge University Press

Lujan, H. L., & Dicarlo, S. E. (2006). Too much teaching, not enough learning: What is the solution? *Advances in Physiology Education, 30*, 17–22.

Machemer P.L., & Crawford, P. (2007). Student perceptions of active learning in a large cross-disciplinary classroom. *Active Learning in Higher Education, 8*(1), 9–30.

Miller, J. E., & MacGilchrist, L. (1996). A model for peer-led work. *Health Education, 2*, 24–29.

Millis, B. J. (2010). Why faculty should adopt cooperative learning approaches. In B. J. Millis (Ed.), *Cooperative Learning in Higher Education: Across the Disciplines, Across the Academy* (pp. 1–9). Sterling, VA: Stylus Publishing.

Millis, B. J., & Cottell, P. G. (1998). *Cooperative learning for higher education faculty*. Phoenix, AZ: American Council on Education and Oryx Press.

Munley, V. G., Garvey, E., & McConnell, M. J. (2010). The Effectiveness of peer tutoring on student achievement at the university level. *American Economic Review: Papers & Proceedings, 100*, 277–282.

Piaget, J. & Cook, M. (1952). *The origins of intelligence in children*. New York, NY: WW Norton & Co.

Pleitz, J., Terry, R. and Judice Campbell, N. (2010). *What they bring with them: Incoming freshmen characteristics and probability of retention*. Poster Presented at CRSDE, November, 1–3, 2010, Mobile, AL.

Rheinheimer, D. C., Grace-Odeleye, B., Francois, G. E., & Kusorgbor, C. (2010). Tutoring: A support strategy for at-risk students. *The Learning Assistance Review, 15*(1), 25–33.

Rubin, D. (1977). Assignment to treatment group on the basis of a covariate. *Journal of Educational Statistics, 2*, 1–26.

Sheets, R. A. (1994). *The effects of training and experience on adult peer tutors in community colleges.* Thesis (Ed.D.), Arizona State University.

Shadish, W. R., Cook, D. T., & Campbell, D. T. (2001). *Experimental and quasi-experimental designs for generalized causal inference.* Stamford, CT: Cengage Learning.

Swengel, E. M. (1991). Cutting education's Gordian knot. *Phi Delta Kappan, 72*(9), 704–710.

Topping, K. J. (1996). The effectiveness of peer tutoring in further and higher education: A typology and review of the literature. *Higher Education, 32*, 321–345.