Journal of College Reading and Learning, 45: 52–66, 2014

ISSN: 1079-0195 print/2332-7413 online
DOI: 10.1080/10790195.2014.949552

Routledge
Taylor & Francis Group

Assessing the Impact of Tutorial Services

CINDY S. TICKNOR, KIMBERLY A. SHAW,
AND TIMOTHY HOWARD
Columbus State University

Many institutions struggle to develop a meaningful way to assess the effectiveness of drop-in tutorial services provided to students. This article discusses the development of a data collection system based on a visitor sign-in system that proved to be an efficient method of gathering assessment data, including frequency of visits, end-of-course grades, and demographic information on student visitors. The data were used to analyze the impact of tutorial services on student grade rates, with special attention given to the effects of service on populations underrepresented (females and Blacks) in science, technology, engineering, and mathematics (STEM) fields. The results showed a significant difference in grade distributions among Black males and provided evidence to support the existence of self-selection biases in the use of tutorial services. The biases may include self-selected use of tutorial services by at-risk students, who are more likely to need support, and by self-motivated students, who are more likely to utilize all available resources to succeed.

KEYWORDS tutoring, tutorial services, mathematics, science, assessment

INTRODUCTION

Most undergraduate institutions offer tutorial services to improve retention and progression rates among students. This is particularly important in science, technology, engineering, and mathematics (STEM) disciplines, in which participation and persistence of female and minority students is consistently low (National Science Foundation, 2013). Various models of tutorial services exist, and current literature suggests several successful strategies that employ supplemental peer instruction. For example, the Emerging Scholars Program (Hsu, Murphy, &

Correspondence concerning this article should be addressed to Cindy S. Ticknor, Schuster Center 124, 4225 University Avenue, Columbus, GA 31907. E-mail: ticknor_cindy@columbusstate.edu

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/ucrl.

Treisman, 2008) introduced calculus workshops promoting collaboration and problem solving, and is known as a seminal model that has been replicated at over 100 institutions (Wilson & Arendale, 2011). Similarly, widely employed Peer-Led Team Learning (PLTL) models have contributed to community building in at-risk groups, increased academic persistence, as well as higher course grades (Amstutz, Wimbush, & Snyder, 2010; Arendale, 2005; Shook & Keup, 2012). Alternatively, supplemental instructional models embed trained peer tutors in courses (Arendale, 2004). Best practices with these models typically include mandatory attendance in peer instructional settings, which is in contrast to free, drop-in tutorial centers that commonly remain part of institutional efforts to improve student performance. As part of our institution's plan to improve student success in introductory STEM courses, we opened a free drop-in tutorial center, the Math and Science Learning Center (MSLC), and developed an affordable assessment system to monitor the effectiveness of peer tutoring services.

The impact of drop-in tutorial centers on retention rates is difficult to assess, and the evaluation is often accomplished through client surveys that tend to focus on characteristics thought to influence the tutoring experiences (e.g., Oates, Paterson, Reilly, & Statham, 2005; Young, 2011). For example, researchers have examined interactions between students and tutors (Evans, Flower, & Holton, 2001; Shaw, Carey, & Mair, 2008), perceptions of tutoring or the roles of tutors (Galbraith & Winterbottom, 2011; Retna, Chong, & Cavana, 2009), and the influence of tutor characteristics such as gender, age, and race on the tutoring experiences (House & Wohlt, 1992). When associated with specific courses, peer tutoring positively impacted student academic performance. For example, optional peer-led supplemental instruction sessions have positively affected end-of-course grades in introductory chemistry courses (Rath, Peterfreund, Bayliss, Runquist, & Simonis, 2012), in anatomy and physiology courses (Hughes, 2011), and in calculus courses (Fayowski & MacMillan, 2008). Similarly, Comfort (2011) found that undergraduates in sports science who attended optional tutoring sessions conducted by peers who were in their final year of the same program had significantly higher grades than those who did not attend. Walker and Dancy (2007) also studied the academic performance of students who utilized a Physics Resource Center for tutorial services for algebra- and calculus-based physics courses. They found significantly lower grades among those who visited the center, but attributed the result to "those who need the help the most being more likely to attend, rather than any adverse effects of tutorial attendance" (p. 138). In other words, students self-selecting to visit tutorial centers may struggle more to understand the material than those who do not visit. Munley, Garvey, and McConnell (2010) studied a drop-in tutoring center with optional peer-led instructional sessions and found that students who were tutored 10 to 20 hours per 14-week semester experienced measurably positive impacts on their grades. When studying the impact of drop-in tutoring on freshman cohorts, Cooper (2010) found higher rates of overall institutional persistence and higher likelihood of good academic standing during their sophomore year. However, the number of visits to the center was not correlated with grades in particular courses. These findings (Comfort, 2011; Cooper, 2010; Fayowski & Macmillan, 2008; Munley et al., 2010; Rath et al., 2012; Walker &

Dancy, 2007) suggest that the use of undergraduates for peer tutoring may be effective when course based, but the impact of drop-in tutorial centers on course grades is still unclear and could be affected by self-selection. Because drop-in centers are commonly used and might be part of institutional support systems that include course-based supplemental instruction, there is a need to develop a strategy for assessing the effectiveness of drop-in centers that serve a variety of courses at a university.

In this article, we discuss the development of our data collection system and a strategy to analyze the impact of tutorial services offered by the MSLC. This analysis considered rates of success in introductory STEM courses by students served versus success rates within the general student population, with special attention on targeted populations that are historically under-represented in STEM fields (those who self-identify as Black and females). Our intent is to provide a model for assessing tutorial services that can be replicated by other institutions, so we will discuss how we developed the system, issues we encountered during the coding and data analysis, and results that we used to recommend strategies to improve assessment of MSLC tutorial services.

The MSLC is a unique institutional resource dedicated to enhancing the learning of math and science through student academic support services and best-practices training for faculty members. The center was created through collaboration between the colleges of Science and Education as part of a STEM initiative, which was initially funded by the University System of Georgia and the Partnership for Reform in Science and Mathematics (PRISM). The academic support services embedded in the MSLC were developed based on two premises: tutorial services improve the content knowledge of both the tutors and the peers they are supporting, and tutorial centers encourage students who are strong in mathematics and science to pursue careers in teaching in the STEM disciplines. Accordingly, the MSLC is directed by a mathematics or science faculty member with experience in education to provide discipline-specific tutor training, mentoring, and advising. Tutor training included initial workshops followed by individualized mentoring and content support, based on observed deficiencies.

In this institutional context, the MSLC tutoring services are free to all students and available on a voluntary drop-in basis, which was deemed to be the most productive mode by which to encourage impromptu student visits, thereby getting students to a tutor at least once. The drop-in aspect of the services was deemed to be necessary in order to improve student perceptions of access to these resources, an issue the organizers felt was important in working with first-generation college students and underrepresented groups. This may differ from other institutions' models, which often require a student who requests tutoring to agree to regular sessions or impose other constraints in exchange for assistance.

DEVELOPING THE DATA COLLECTION SYSTEM

In early 2009, we began developing a data collection system to determine the effectiveness of the tutoring services offered at the MSLC. This system was implemented in June 2009. Students visiting the MSLC signed in at the entrance of the center on a laptop computer that recorded the date and time, and also queried the reason for their visit. Possible reasons for visiting included "tutoring," "quiet area for study," "computer use," and "other reasons;" multiple reasons could be selected. If the students indicated tutoring as a reason for their visit, they were prompted to choose from the courses in their course schedule that term. First-time visitors to the MSLC were offered an electronic informed consent document, and their participation in this evaluation study was requested. This program was designed locally and is completely customizable, avoiding the cost of software and card swipe systems used at other institutions (Cooper, 2010). Such software may be preferable if scheduling features are useful to institutions, but we needed a means of collecting Institutional Review Board-approved data, only, not the scheduling features of available software.

The database was designed to produce two reports: a usage report to track the number and types of visits to the MSLC, and a summary report for use at the completion of each academic term. Usage reports are designed to help match tutor schedules with student demand. In addition to using the data to inform long-term resource planning for the MSLC, the usage report could also be used to inform math and science faculty about the percentage of their students requesting tutoring services at the MSLC. Summary reports are created after the end of the semester and include the total number of visits for each student as well as each visitor's course grades earned during the semester. The usage and summary reports utilized student institutional identification numbers that allowed us to collect demographic information about the students from our Office of Institutional Research.

Collaboration between the MSLC director, university programmers, and institutional research personnel was essential in the development of this system. Conversations among these personnel reduced impediments to data collection and greatly simplified the process. However, there are still three known flaws in the data-acquisition program. First, not all visits are recorded. Visitors do not always sign in, even with the computer readily available in the entrance and staff actively encouraging visitors to do so. Second, student clients do not always remember their institutional identification numbers, in which case some of their demographic and performance data cannot be retrieved. Third, changes in visit foci are unlikely to be counted. For example, if a student initially designated only "quiet place to study" as a reason for the visit, but then decided to ask a tutor for assistance, the report would not reflect that tutoring support was provided. All of these flaws contribute to an undercounting of students served by this program.

CODING THE DATA

After developing the data collection system and producing a usage data report, we retrieved an *institutional* data report from our Office of Institutional Research for the academic year 2010–2011; the report included the student identification numbers of all students enrolled in introductory science and mathematics courses. For each student enrolled, we collected the grades earned in each science or math course, demographic information, institutional grade point average (GPA), and standardized test scores (SAT and/or ACT). Although SAT and ACT scores may not provide a perfect means of comparing student ability, previous studies at the institution (Shaw, Gurkas, & Webster, 2012) indicated that SAT Math scores (and ACT Math scores) tend to correlate with student success in introductory STEM courses. All subjects were assigned a variable called "Test Score" that was either their ACT Math score or an SAT Math score converted using Dorans's (1999) concordance table. Although the College Board has published more recent concordance tables (ACT and SAT Concordance Tables, 2009), they include only ACT Composite scores in concordance with the sum of SAT Math and Critical Reading scores. Students who had neither SAT nor ACT scores on record with the university were excluded from the analysis. To verify consistency, we conducted a linear regression of students who had reported both SAT Math and ACT Math scores, in order to determine if our assigned "Test Score," which was generated from the SAT Math score, was correlated with the ACT Math score. We found a strong positive correlation $(R^2 = 0.68, p = 0.01)$ between actual ACT Math scores and the variable "Test Score."

Next, we merged information from the data in the institutional data report and the usage data report; in many ways, this was the most challenging aspect of our work. We initially matched only student identification numbers in each list and coded students as "visitors," a fairly simple procedure because the identification numbers were simple to match. First, some students had received tutoring for more than one course, and others were enrolled in more than one math or science course, yet requested tutoring in only one of those courses. Second, naming conventions for courses in the two data reports differed, requiring extensive work in order to match the data appropriately. Those who hope to replicate the system should address these issues in the development phase so that all reports can be easily matched.

After aligning all course names, we were able to match courses identified by "visitors" in the usage data report (i.e., the courses for which they received tutoring) with courses in the institutional data report and were able to identify and code courses and course grades as "served." Thus, we were able to identify students as "visitor" or "nonvisitor" to the center and the courses as "served" or "not served." For example, a student might have taken both biology and calculus in the same semester, but sought tutoring only for calculus. In such a case, we included that student in the analysis as a visitor, but included that student as "served" only for calculus. Once the data were coded, we developed the following research questions:

- Do students who receive tutorial services for a course (i.e., "served" students) earn better grades than those who do not visit the MSLC?
- Do different target populations who receive tutorial services for a course earn different grades than nonvisitors from the same background?
- Did either the number of times that a student visited the MSLC or the date of the student's first visit impact course performance?
- Are students who are considered underrepresented groups within the STEM disciplines utilizing the MSLC?
- After reviewing our initial analysis of the data, what appeared to be a self-selection in our data stimulated an additional question: Are students who are earning A's and B's less likely to visit the MSLC?

The target populations for the research were female and Black students, as self-reported to our institutional demographic database.

RESULTS OF INITIAL ANALYSES

We addressed the first question by comparing the grade distributions of students who never logged into the MSLC and those who were served by the MSLC, restricting our dataset to courses that had been identified by at least five students who entered the MSLC for assistance. By running a cross tabulation and a Chi-squared test for independence, we determined that those served by the MSLC had significantly different grade distributions, X^2 (4, N = 2828) = 24.10, $p < 0.001$, with a small effect size of Cramer's V = 0.09. In a second iteration of the analysis, we removed all cases of incomplete grades and those in which students earned the grade "W" for withdrawal. Students must withdraw from a course by the end of the fourth week of class in order to earn a "W" and would be unlikely to benefit from visiting the MSLC. Indeed, only four students who visited the MSLC earned a "W" or an "I," compared with 112 students who did not use the MSLC. The analysis again found significant differences in the distributions X^2(4, N = 2712) = 20.15, $p < 0.001$, with Cramer's V = 0.089 indicating a small effect size (Table 1).

Because the significant finding of the Chi-square analysis might be influenced by the large size of the sample, we also randomly selected 20% of the sample to verify the results, finding that X^2 (4, N = 542) = 9.4, $p = 0.052$, with Cramer's V = 0.13 again indicating a small effect. Because the average Test Score for all categories appeared to be lower for students served by the MSLC, we also attempted two additional regression analyses. The first was a multiple regression that controlled for the covariant, Test Score, which found a significant correlation between Test Scores and course grades, $\beta = 0.297$, $t = 16.28$, $p < .001$, but whether a student was served at the MSLC was not a significant predictor in of the course grade ($\beta = 0.029$, t =1.62, $p < 0.11$). Because multiple regression analysis treated grades as continuous rather than ordinal we again tested the data using an ordinal logit regression and found that a significant relationship did not exist between being served at the MSLC and Test Scores.

We repeated the analysis considering two additional variables: the number of visits per semester and the date of the first visit to the MSLC. Using the prior dataset, which was restricted to courses with more than five visitors and those who

TABLE 1. Cross-Tabulation of Nonvisitors vs. Served Students at the MSLC

	Grades						
	A	B	C	D	F/WF	X^2	V
Nonvisitors	425 (416)	563 (547)	608 (615)	294 (318)	461 (454)	20.15*	0.09
Mean Test Score:	*23.2*	*21.7*	*21.0*	*19.7*	*21.2*		
Served	55 (64)	68 (84)	102 (95)	73 (49)**	63 (70)		
Mean Test Score:	*21.4*	*19.6*	*19.0*	*18.9*	*18.1*		

Notes. Expected values appear in parentheses. * = $p \leq 0.001$. **Largest contributor to the X^2 statistic.

TABLE 2. Cross-Tabulation of Nonvisitors vs. Students Served at the MSLC More Than One Time

Grades	A	B	C	D	F/WF	X^2	V
Nonvisitors:	425 (416)	563 (547)	608 (615)	294 (318)	461 (454)	19.4*	0.09
Mean Test Score:	*23.2*	*21.7*	*21.0*	*19.7*	*19.7*		
Served:	43 (49)	52 (65)	79 (72)	58 (37)**	44 (53)		
Mean Test Score:	*20.9*	*19.1*	*19.3*	*19.1*	*18.1*		

Notes. Expected values appear in parentheses. * = $p \leq 0.001$. **Largest contributor to the X^2 statistic.

did not earn a "W" or an "I," we additionally excluded all cases in which students visited the MSLC only one time. The analysis again found significant differences in the distributions X^2 (4, N = 2627) = 19.42, $p < 0.001$, with Cramer's V = 0.086 indicating a small effect size (Table 2). We repeated the process, excluding students who visited fewer than five times, and found consistent results, with a significant difference in distributions, X^2(4, N = 2507) = 17.25, $p < 0.002$, with Cramer's V = 0.083 indicating a small effect size. Therefore, even when we restricted our analysis to the most frequent users of the MSLC, we observed the same significant difference in grade distributions.

Next we utilized the time-stamp feature of the assessment system to calculate the date of first use by a student, quantified as the proportion of the semester that had passed prior to the first visit. We wanted to determine if there was a correlation between the end-of-course grade, the number of visits to the center, and the date of the first visit. We conducted a multiple regression analysis of all those served by the MSLC and did not find a significant correlation. We examined the grade distributions considering those who had first visited the center during the first half of the semester compared to those who visited in the last half of the semester, and did not find a difference in the distribution that was statistically significant (Table 3). We did note that of the 102 students who came to the MSLC during the last half of the term, 45 (44%) came only one time, and 73 (72%) came only twice. Twenty nine of the students came only one time during the last quarter of the semester.

After considering the impact on the general population of students who visited the MSLC, we also wanted to determine if similar effects were evident in our target population. Using Chi-square analyses on each group, the only significant

TABLE 3. Cross-Tabulation of Students First Served During the First Half vs. the Last Half of the Semester

Grades	A	B	C	D	F/WF	X^2	V
First Half	46 (39)	50 (49)	66 (72)	50 (52)	47 (45)	7.07	0.13
Mean Test Score:	*21.5*	*19.2*	*18.8*	*18.9*	*18.3*		
Second Half	9 (16)	18 (19)	36 (29)	23 (21)	16 (18)		
Mean Test Score:	*20.9*	*20.7*	*19.4*	*18.8*	*17.4*		

Note. X^2 statistic was not significant. Expected values appear in parentheses.

TABLE 4. Cross-Tabulation of Nonvisitors vs. Served Black Males at the MSLC

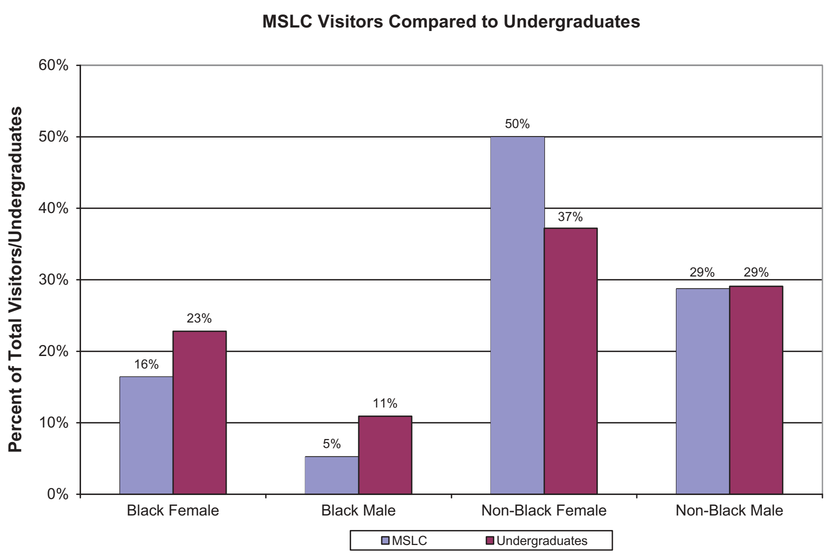
	Grades						
	A	B	C	D	F/WF	X^2	V
Nonvisitors	28(26)	47(41)	69(75)	30(36)	56(52)	12.79*	0.21
Mean Test Score:	*21.6*	*20.3*	*20.4*	*19.1*	*19.6*		
Served	5(7)	6(12)	27(21)	16(10)**	10(14)		
Mean Test Score:	*21.0*	*19.2*	*18.7*	*17.4*	*18.8*		

Note: * $= p \leq 0.05$. **Largest contributor to the X^2 statistic. Expected values appear in parentheses.

differences noted were in the Black males group, which showed a significant difference in grade distributions with $X^2(4, \text{N} = 294) = 12.79$, $p < 0.012$, and Cramer's V = 0.21 indicating moderate effect size (Table 4). We then examined the data, as we did with the general population, restricting the population to those who visited the MSLC more than one time. Unfortunately, we did not find a significant relationship for Black students (N = 101) or women (N = 178), and there was not a sufficient number of Black males to conduct a meaningful analysis (N = 28).

Overall, it appeared that the MSLC was being underutilized by some of our many target populations. Based on our usage data, 66% of the visitors to the MSLC were female, yet females comprise only 60% of our undergraduate population. As shown in Figure 1, the distribution of the MSLC visitors compared to the total undergraduate population indicate that the center is being visited by a

FIGURE 1. Distribution of MSLC visitors compared to undergraduates.



larger proportion of female students who did not self-identify as Black (50% of the MSLC visitors vs. 37% of the overall undergraduate population). However, Black students did not use the MSLC proportionate to our overall undergraduate population distribution. Overall, 1,110 students visited the MSLC during the academic year of this study, including 182 (16%) Black females and 58 (5%) Black males, compared to institutional enrollments of 23% and 11%, respectively.

Discussion: Initial Results

The course grade distribution of students who reported using the MSLC for a particular math or science course varied significantly from those who did not visit the MSLC. The grade distribution of those served by the MSLC was characterized by a larger-than-expected frequency of C's and D's, with a smaller-than-expected number of A's, B's, and F's. The greatest contributor to the X^2 (20.15) statistic was the increase in the number of D's (11.94) and decrease in the number of B's (3.05). When we restricted our analysis to students who used the MSLC at least once, we found similar patterns in the grade distributions. The greatest contributor to the X^2 statistics was the greater-than-expected number of D's (11.9). The next two largest contributors were fewer B's (2.46) and fewer F's than expected (1.55). Although significant patterns were not noted in White females or Black females, Black male students who reported using the MSLC for a particular course had grade distributions that were significantly different than those of their peers who did not use the center. The grade distribution of Black males served by the MSLC was similar in character to all those served by the MSLC; the distribution had a larger-than-expected frequency of C's and D's, with a smaller-than-expected number of A's, B's, and F's. The greatest contributor to the X^2 (12.80) statistic was the increase in the number of D's (3.58) and decrease in the number of B's (2.66). This would imply positive impact on course performance, but from a group that tends to underutilize the MSLC. Interestingly, in both overall and Black male grade distributions, the average Test Score of those who used the MSLC was consistently below those who did not, in every grade category. However, when controlling for the effect of Test Score, visiting the MSLC was not a statistically significant predictor of grades in courses in multiple or ordinal logit regression analyses. One possible explanation for the lack of correlation might be due to a decrease in linearity of the data, because those who are served by the MSLC have fewer A's, B's, and F's but more C's and D's. Therefore, examining grade distributions, along with course GPA, proved informative.

Results of Second Analysis Examining Self-Selection

The pattern in the grade distribution for all users and Black male users of MSLC invited consideration of two self-selection biases that could be impacting the results. First, students who are doing well in class might be less likely to visit the center, which prompted an additional research question: "Are students who are earning A's and B's less likely to visit the MSLC?" This type of effect was suggested by Walker and Dancy (2007) when they observed significantly lower grades for those who visited a drop-in tutorial center. Based on the above distributions, this may be indicated because 41% of students in the courses analyzed earned A's and B's, but only 34% of those using the MSLC earned A's and B's. A second potential case of self-selection could be that students who choose to visit the MSLC for tutorial assistance are more motivated to succeed and would perhaps be more likely to outperform their peers with or without the services of the MSLC. MSLC usage data was revisited to address the question, "Are students who visit the MSLC more likely to do well in their math and science courses, regardless of receiving tutoring?"

Using the same restrictions on our dataset, the question was addressed by calculating average grade performance in introductory courses for three groups of individuals: (a) students who did not visit the MSLC, (b) students who visited the MSLC but did not receive tutoring for the course, and (c) students who visited the MSLC and received tutoring for the course (see Table 5). For example, if a student visited the MSLC and received tutoring for college algebra but not chemistry, their grade performance would be in group "c" with respect to college algebra and group "b" with respect to chemistry. A one-way, between-subjects analysis of variance (ANOVA) was then conducted to determine if there was a significant difference in course performance. There was a significant effect at the $p < 0.001$ for the three groups [F(2, 3136) = 14.65, $p = 0.000$]. Tukey Honest Significant Difference post-hoc comparisons of the three groups indicated significantly greater course GPAs among (b) Visitors, Courses Not Tutored (M = 2.42, 95% CI [2.28, 2.56]) and (a) Nonvisitors (M = 1.99, 95% CI [1.93, 2.04], $p = 0.000$), as well as with (c) Visitors, Courses Tutored (M = 1.92, 95% CI [1.79, 2.06], $p = 0.000$; Table 5).

We repeated the analysis, restricting our cases to those who visited the MSLC more than one time and again found a significant effect for the three groups [F = (2, 2910) = 4.663, $p = 010$]. The post-hoc comparison found significantly greater course GPAs among (b) Visitors, Courses Not Tutored (M = 2.32, 95% CI [2.12, 2.56]) and (a) Nonvisitors (M = 1.99, 95% CI [1.93, 2.04], $p = 0.008$) and (c) Visitors, Courses Tutored (M = 1.95, 95% CI [1.80, 2.10], $p = 0.018$; Table 6). We also compared each group's Test Scores and found a statistically significant effect at the $p < 0.001$ level F[(2, 2910) = 29.64, $p = .000$] with a post-hoc analysis revealing significant differences at $p < 0.05$ among all three mean test

TABLE 5. ANOVA of Nonvisitors & Visitors at the MSLC

Factors:	*N*	*Mean* (*s*)	*F*	*95% CI*
(a) Nonvisitors $_{b**,}$	2463	1.99 (1.40)	14.65**	[1.93, 2.04]
(b) Visitors, Courses Not Tutored $_{a**,c**}$	311	2.42 (1.29)		[2.28, 2.56]
(c) Visitors, Courses Tutored$_{,b**}$	365	1.92 (1.31)		[1.79, 2.06]

Note: ** = $p \leq 0.001$ Standard deviations in parentheses. Subscripts indicate significant differences found among factors.

TABLE 6. ANOVA of Nonvisitors & Those Who Visited the MSLC at Least Twice

Factors:	*N*	*Mean (s)*	*F*	*95% CI*
(a) Nonvisitors $_{b*}$	2463	1.99 (1.40)	4.66**	[1.93, 2.04]
(b) Visitors, Courses Not Tutored $_{a*,c**}$	171	2.32 (1.32)		[2.12, 2.51]
(c) Visitors, Courses Tutored $_{b**}$	279	1.95 (1.30)		[1.80, 2.10]

Note: * $= p \leq 0.05$, ** $= p \leq 0.01$ Standard deviations in parentheses. Subscripts indicate significant differences found among factors.

scores. Nonvisitors had the highest mean [M = 21.1, 95% CI [20.9, 21.2], followed by (b) Visitors, Courses Not Tutored [M = 20.3, 95% CI [17.8, 20.9], and (c) Visitors, Courses Tutored [M = 19.3, 95% CI [18.9, 19.7]. This provides some evidence that students who visited the center were more likely to perform better in math and science courses, despite having lower standardized test scores than nonvisitors ($p = 0.027$), and it could indicate that they were more motivated to succeed or take action to improve their course performance. In addition, those who did receive tutoring had significantly lower standardized test scores than nonvisitors ($p = 0.017$) yet did not have significantly different averages on their end-of-course grades.

Discussion: The Case for Self-Selection

By employing an analysis that examines student grade distributions, evidence emerged that supports the case for self-selection bias. First, the smaller proportion of A's and B's in the courses served may be a result of the self-selection caused by students arriving at the center when they felt they are underperforming in a course. However, the smaller proportion of failing grades (F and WF) and greater proportion of D's in this subgroup may indicate the MSLC's ability to improve the expected grades of students. This appears to be a likely trend, particularly with Black male students. Therefore, this self-selection bias would cause a shift in course grade distributions but would not necessarily impact the end-of-course grade average. In other words, we were able to observe grades tending toward C's and Ds' compared to those of nonvisitors, although, at the same time, we did not find a significant difference in end-of-course averages. This might help to explain why regression analyses often do not find significant differences in course performance for those who receive drop-in tutorial services. Multiple regression analyses such as ours assume linear relationships exist between independent and dependent variables, and the observed shift in grade distribution might cause an inaccurate estimation of the true effect of tutorial services. Second, by examining the end-of-course grades of visitors and nonvisitors, we were able to see that students who visited the center outperformed nonvisitors in courses in which they did *not* receive tutoring. This provides evidence that visitors to the center might have other personal attributes that contribute to their success, including motivational factors.

Alternatively, the learning strategies, instructional methods, or positive reinforcement received during their visit to the MSLC might have been transferable to their other studies.

One way to understand these differences in behavior could be in terms of a student's readiness for change. Prochaska, DiClemente, and Norcross (1992) developed a five-stage model delineating a general process of change, and Weinsheimer (1998) argued that these stages could be employed to frame our understanding of students who seek support from tutorial services. If students are not yet aware of their level of academic support need (precontemplation stage), considering taking action (contemplation stage), or just planning to use tutorial services (planning stage), they are not likely to productively use services, even if required to do so. Once they reach the action stage, students can expend great time and energy to change their academic success, and during the maintenance stage they might work to avoid relapse. Using this perspective, it could be that those who utilized tutorial services were ready to change their academic performances and were more motivated to succeed. Conversely, those who were not yet ready to change their study habits were less likely to seek help and less likely to utilize help effectively or consistently (Weinsheimer, 1998). In turn, this supports the claim that there could be a self-selection bias underlying the statistics. In other words, students who seek support services are more likely to improve their grades, not necessarily from the services provided by the center, but because they are (in part) motivated to succeed and are cognitively ready to benefit from help provided by tutoring centers. Questions of motivation or readiness can be answered fully only by collecting additional qualitative data outside the scope of this study.

SUMMARY AND FUTURE CONSIDERATIONS

We believe that using the current analysis of grade distributions as a base point of reference will give us the foundational understanding of how biases might affect the data and explain the varying results of drop-in tutoring and course performance (Cooper, 2010; Walker & Dancy, 2007). We are especially intrigued by the increased proportion of D's in the grade distributions of those who are served by the MSLC, and it could be that students are requesting tutorial services too late in the semester for the effort to significantly impact their grades. Indeed, a large proportion of students who visited the MSLC only one time came during the last quarter of each semester. With respect to our targeted populations, our results suggest that drop-in tutoring may be underutilized by Black students, which is of concern because males appeared to benefit from tutorial services.

Because of our ability to capture and analyze usage data, we developed specific recommendations for the MSLC. First, data suggested peak tutoring demand times, which could inform more efficient tutor scheduling. Second, the underutilization by Black students should be more extensively researched. The MSLC was designed for drop-in tutorial services, which rely on students seeking assistance, yet that action is likely impeded by the students' readiness to change. The decision of organizational structure for academic support, such as drop-in tutoring, typically has the greatest impact on achievement (Weinsheimer, 1998). Therefore, the MSLC might need to adjust strategies to encourage use by students who are contemplating seeking assistance but are dissuaded by perceptions that receiving tutoring means they are academically deficient. For example, the MSLC should perhaps promote activities that focus on course-related learning, which would address the difficult nature of the material, as opposed to deficiencies in student ability. Tutors would require training not only in recognizing deficient skills, but also in providing socially supportive environments. In addition, the MSLC could encourage use earlier in each semester by promoting the center for quiet studying, coordinating study groups, or conducting study skills workshops. The latter might focus on overcoming impediments to seeking assistance in math and science learning for multiple audiences: tutors, faculty, and students in general.

As these recommendations are adopted at our institution, additional research projects might be embedded within the assessment system to evaluate the effectiveness of the new strategies. Currently, the system gathers information on students' reasons for visiting the MSLC, but it could be expanded to capture students' stages of readiness. This is a distinct advantage of creating a data collection system, rather than using scheduling software. In addition, we are planning to add an exit survey that records the duration of students' visits and their perceptions of the services they received. This will permit future studies to analyze whether duration or quality of tutoring visits impacts course performance. Finally, our ability to assess drop-in services is vital as we develop supplemental instruction programs based on best practices. We are currently collecting data that will allow us to compare drop-in services with an alternative peer instructional model, and should inform recommendations for the allocation of future resources designed to improve student success.

Overall, the development of the assessment system and generation of the MSLC usage reports have allowed us not only to establish baseline data that can be used to evaluate the impact of tutorial services but also to discover important insights for conducting the data analysis. We believe it is critical in any future analyses to include an examination of grade distributions in conjunction with regression analyses because self-selection biases may cause variation in grade distributions but not overall averages. We believe our assessment system and analytical approach could be easily replicated at other institutions and can be customized based on the organizational structure and goals of tutorial centers.

ABOUT THE AUTHORS

Cindy S. Ticknor, PhD, is a Professor of Mathematics Education and Interim Dean of the Honors College at Columbus State University. She conducts research on the impact of high impact practices and support services on student retention and progression.

Kimberly A. Shaw, PhD, is a Professor of Physics and Co-Director of the UTeach Program at Columbus State University. Her research interests include factors that impact student learning and success in college STEM settings.

Timothy Howard, PhD, is a Professor of Mathematics and Director of the Math & Science Learning Center at Columbus State University. His research interests include collegiate teaching and learning, factors affecting college student success, mathematical puzzles and games, and the preparation of K-12 teachers.

REFERENCES

ACT and SAT Concordance Tables. (2009). Retrieved from http://research.collegeboard.org/sites/default/files/publications/2012/7/researchnote-2009-40-act-sat-concordance-tables.pdf

Amstutz, M., Wimbush, K., & Snyder, D. (2010). Effectiveness and student demographics of peer-led study groups in undergraduate animal science courses. *NACTA Journal*, *54*(1), 76–81.

Arendale, D. R. (2004). Pathways of persistence: A review of postsecondary peer cooperative learning programs. In I. M. Duranczyk, J. L. Higbee, & D. B. Lundell (Eds.), *Best practices for access and retention in higher education* (pp. 27–42). Retrieved from: http://www.cehd.umn.edu/crdeul/pdf/monograph/5-a.pdf

Arendale, D. R. (2005). *Postsecondary peer cooperative learning programs: Annotated bibliography*. Retrieved from: http://www.eric.ed.gov/contentdelivery/servlet/ERICServlet?accno=ED489957

Comfort, P. (2011). The effect of peer tutoring on academic achievement during practical assessments in applied sports science students. *Innovations in Education & Teaching International*, *48*(2), 207–211.

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

Dorans, N. J. (1999). *Correspondences between ACT™ and SAT® I scores* (College Board Report No. 99-1 ETS RR No. 99-2). New York, NY: College Entrance Examination Board.

Evans, W., Flower, J., & Holton, D. (2001). Peer tutoring in first-year undergraduate mathematics. *International Journal of Mathematical Education in Science and Technology*, *32*(2), 161–173.

Fayowski, V., & MacMillan, P. (2008). An evaluation of the supplemental instruction programme in a first year calculus course. *International Journal of Mathematical Education in Science & Technology*, *39*(7), 843–855.

Galbraith, J., & Winterbottom, M. (2011). Peer-tutoring: What's in it for the tutor? *Educational Studies*, *37*(3), 321–332.

House, J. D., & Wohlt, V. (1992). Tutoring outcomes of academically underprepared adolescent minority students as a function of student and tutor characteristics. *Journal of Genetic Psychology*, *153*(2), 225–227.

Hsu, E., Murphy, T. J., & Treisman, U. (2008). Supporting high achievement in introductory mathematics courses: What we have learned from 30 years of the emerging scholars program. In M. Carlson & C. Rasmussen (Eds.), *Making the connection: Research and teaching in undergraduate mathematics education* (pp. 205–220). Washington, DC: Mathematical Association of America.

Hughes, K. S. (2011). Peer-assisted learning strategies in human anatomy & physiology. *American Biology Teacher*, *73*(3), 144–147.

Munley, V., Garvey, E., & McConnell, M. (2010). The effectiveness of peer tutoring on student achievement at the university level. *American Economic Review: Papers & Proceedings 100*(2), 277–282.

National Science Foundation, National Center for Science and Engineering Statistics. (2013). *Women, minorities, and persons with disabilities in science and engineering: 2013* (Special Report NSF 13-304). Arlington, VA. Retrieved from http://www.nsf.gov/statistics/wmpd/

Oates, G., Paterson, J., Reilly, I., & Statham, M. (2005). Effective tutorial programmes in tertiary mathematics. *International Journal of Mathematical Education in Science and Technology*, *36*(7), 731–739.

Prochaska, J., DiClemente, C., & Norcross, J. (1992). In search of how people change. *American Psychologist*, *47*(9), 1102–1114.

Rath, K., Peterfreund, A., Bayliss, F., Runquist, E., & Simonis, U. (2012). Impact of supplemental instruction in entry-level chemistry courses at a midsized public university. *Journal of Chemical Education*, *89*(4), 449–455.

Retna, K. S., Chong, E., & Cavana, R. Y. (2009). Tutors and tutorial: Students' perceptions in a New Zealand university. *Journal of Higher Education Policy & Management*, *31*(3), 251–260.

Shaw, K. A., Gurkas, P., & Webster, Z. T. (2012). An analysis of factors expected to impact end-of-course grades in introductory science classes. *Perspectives in Learning*, *13*(1), 4–16.
Shaw, L., Carey, P., & Mair, M. (2008). Studying interaction in undergraduate tutorials: Results from a small-scale evaluation. *Teaching in Higher Education*, *13*(6), 703–714.
Shook, J. L., & Keup, J. R. (2012). The benefits of peer leader programs: An overview from the literature. *New Directions for Higher Education*, *157*, 5–16.
Walker, K. N., & Dancy, M. H. (2007). Investigation and evaluation of a physics tutorial center. *AIP Conference Proceedings*, *883*(1), 137–140.
Weinsheimer, J. (1998). *Providing effective tutorial services*. Washington, DC: National TRIO Clearinghouse, National Council of Educational Opportunity Association.
Wilson, W. R., & Arendale, D. (2011). Peer educators in learning assistance programs: Best practices for new programs. *New Directions for Student Services*, *133*, 41–53.
Young, E. (2011). Onsite peer tutoring in mathematics content courses for pre-service teachers. *Issues in the Undergraduate Mathematics Preparation of School Teachers*, *3*, 1–8.