

Article

Increasing Persistence in Undergraduate Science Majors: A Model for Institutional Support of Underrepresented Students

Brit Toven-Lindsey,* Marc Levis-Fitzgerald,* Paul H. Barber,[†] and Tama Hasson[‡]

*Office of Instructional Development, [†]Ecology and Evolutionary Biology, and [‡]Division of Undergraduate Education, University of California, Los Angeles, Los Angeles, CA 90095

Submitted May 5, 2014; Revised January 29, 2015; Accepted February 2, 2015
Monitoring Editor: Alison Gammie

The 6-yr degree-completion rate of undergraduate science, technology, engineering, and mathematics (STEM) majors at U.S. colleges and universities is less than 40%. Persistence among women and underrepresented minorities (URMs), including African-American, Latino/a, Native American, and Pacific Islander students, is even more troubling, as these students leave STEM majors at significantly higher rates than their non-URM peers. This study utilizes a matched comparison group design to examine the academic achievement and persistence of students enrolled in the Program for Excellence in Education and Research in the Sciences (PEERS), an academic support program at the University of California, Los Angeles, for first- and second-year science majors from underrepresented backgrounds. Results indicate that PEERS students, on average, earned higher grades in most “gatekeeper” chemistry and math courses, had a higher cumulative grade point average, completed more science courses, and persisted in a science major at significantly higher rates than the comparison group. With its holistic approach focused on academics, counseling, creating a supportive community, and exposure to research, the PEERS program serves as an excellent model for universities interested in and committed to improving persistence of underrepresented science majors and closing the achievement gap.

INTRODUCTION

The number of science, technology, engineering, and mathematics (STEM) majors as an overall percentage of the undergraduate population has fallen substantially over the past several decades (National Board of Sciences, 2010). A major factor driving this decline is the failure of entering STEM majors to complete their degrees. Nationwide, 6-yr degree-completion rates in STEM majors are less than 40%

(President’s Council of Advisors on Science and Technology [PCAST], 2012), raising concerns for the ability of the United States to maintain its competitiveness in science and technology fields (Hira, 2010). Low rates of STEM persistence are particularly troubling among women and underrepresented minority (URM) students (URM students include African-American, Latino/a, Chicano/a, Native American, and Pacific Islander students). While women and URM students account for nearly 70% of college enrollment, they are underrepresented among STEM degree holders, because they leave STEM majors at substantially higher rates than their non-URM male peers (PCAST, 2012).

URM students entering U.S. colleges are just as likely as their non-URM peers to aspire to complete a STEM major (Crisp *et al.*, 2009; Koenig, 2009; Hurtado *et al.*, 2010). Yet in 2009, 37.5% of white and Asian-American students completed their STEM degrees after 5 yr, while the average completion rates for black/African-American, Latino/a, and Native American students were 22.1, 18.4 and 18.8%, respectively (Hurtado *et al.*, 2010). This persistence gap results in fewer URM students entering the STEM workforce.

CBE Life Sci Educ June 1, 2015 14:ar12

DOI:10.1187/cbe.14-05-0082

Address correspondence to: Tama Hasson (thasson@college.ucla.edu).

© 2015 B. Toven-Lindsey *et al.* CBE—Life Sciences Education © 2015 The American Society for Cell Biology. This article is distributed by The American Society for Cell Biology under license from the author(s). It is available to the public under an Attribution–Non-commercial–Share Alike 3.0 Unported Creative Commons License (<http://creativecommons.org/licenses/by-nc-sa/3.0>).

“ASCB®” and “The American Society for Cell Biology ®” are registered trademarks of The American Society for Cell Biology.

Currently, blacks, Hispanics, and Native Americans account for only 7.1% of the biological, biomedical, and life sciences workforce, despite accounting for 27.9% of the total U.S. population (National Academy of Sciences, National Academy of Engineering, and Institute of Medicine, 2011; National Science Foundation [NSF], 2012, Table 9-6).

Factors That Contribute to URM Persistence in STEM

Multiple factors contribute to the high attrition of underrepresented students in STEM majors. Socially, many of these students face challenges transitioning to college (Cooper *et al.*, 2005; Museus and Quayle, 2009), in part because they are more likely to be first-generation college students (Terenzini *et al.*, 1996; Choy *et al.*, 2000; McCarron and Inkelas, 2006). These challenges can be further exacerbated by perceptions of an unwelcoming academic culture in science and math departments (Ong *et al.*, 2011; Beasley and Fischer, 2012). Academically, many students struggle to complete introductory science and math courses based on insufficient preparation in high school (Chang *et al.*, 2014; Elliott *et al.*, 1996) and challenges staying engaged in large lecture-style courses with limited opportunities for interaction with professors (Labov, 2004; Johnson, 2007; Gasiewski *et al.*, 2012).

On the basis of social and cognitive psychology, Graham *et al.* (2013) proposed a persistence framework focused on increasing confidence and motivation of STEM majors. They highlighted early research experience, active learning in introductory courses, and learning communities as critical components for effective learning and feeling like a scientist. Indeed, academic support programs that employ aspects of this model and aim to support underrepresented students in STEM fields, such as the Meyerhoff Scholars Program at the University of Maryland–Baltimore County (Stolle-McAllister *et al.*, 2011) and the Biology Scholars Program at the University of California, Berkeley (Matsui *et al.*, 2003), can have a major impact, increasing persistence rates of URM students in STEM majors to levels two to four times the national average (Summers and Hrabowski, 2006).

Studies of STEM academic support programs typically measure program effectiveness by comparing the persistence rates of participants with university or national averages (e.g., White *et al.*, 2008). While persistence is a critical metric, it can be difficult to determine the factors contributing to STEM retention. Indeed, few studies employ control groups to more fully explore the variables promoting persistence of underrepresented students in undergraduate STEM majors. This study utilizes a matched comparison group design and propensity score matching to examine the academic achievement and persistence of undergraduate students enrolled in an academic support program for STEM majors at the University of California, Los Angeles (UCLA). Specifically, we use a variety of metrics of academic achievement to determine whether students in the Program for Excellence in Education and Research in the Sciences (PEERS) achieve the same level of academic success as non-PEERS students, and how this achievement impacts their commitment to and persistence in STEM majors.

URM Persistence at UCLA

UCLA is a highly selective research university in Southern California. The overall 5-yr degree-completion rate for STEM

students at UCLA is 65%, much higher than the national average. Yet significant disparities exist between URM and non-URM students. For example, registrar data show that, for students entering UCLA with a declared STEM major between 2004 and 2006, nearly 70% of non-URM students completed their STEM degree in 5 yr, while the degree-completion rate of URM students in STEM was only 39%.

As in many U.S. universities, life and physical sciences degree programs at UCLA are characterized by a highly regimented curriculum of core introductory math and science courses. Successful completion of these “gatekeeper” courses during the first 2 yr enables students to move on to upper-division courses and ultimately graduate in a timely manner. However, these same math and science courses can be a major impediment to persistence of underrepresented students in STEM majors (Fries-Britt *et al.*, 2010; Gasiewski *et al.*, 2012). As confidence and motivation are critical for STEM persistence (Graham *et al.*, 2013), it follows that increasing success in the introductory core science curriculum is an essential component of encouraging higher levels of persistence among STEM majors.

PEERS is a 2-yr, cohort-based academic support program designed to promote persistence of students from disadvantaged backgrounds in life and physical sciences majors at UCLA. In this study, we compare the academic performance of students in the PEERS program with a matched comparison group of non-PEERS students to assess the impact of PEERS on academic success and persistence in science majors during the first 2 yr at UCLA. Specifically, we examine students’ 1) grades in select lower-division chemistry and math courses, 2) total number of science courses completed, 3) academic performance as measured by cumulative grade point average (GPA), and 4) persistence in science and math majors.

METHODS AND RESULTS

The PEERS Program

The PEERS program at UCLA was established in 2003 to address the discrepancy between success and persistence of life and physical sciences majors from underrepresented backgrounds. Based on research related to students’ first-year experiences (Barefoot, 2000; Tinto, 2005; Reason *et al.*, 2006), programmatic elements include: 1) academic and career seminars; 2) holistic academic counseling; 3) research seminars; and 4) Treisman-style collaborative-learning workshops for the first year of math, chemistry, and physics courses (Treisman, 1992). Combined, these activities provide students with encouragement, academic preparation, and positive peer-group motivation, factors shown to encourage persistence in science and math majors (Bonous-Hammarth, 2000; Barlow and Villarejo, 2004; Peterfreund *et al.*, 2007; Chang *et al.*, 2011; Walton and Cohen, 2011). Additionally, PEERS socializes students to the roles and expectations of the institution and their academic major, factors that are also positively correlated with persistence (Chang *et al.*, 2008).

Defining a Control Group: Three Approaches

Matched Comparison Group. Following the methods of DeAngelo and Hasson (2009), we used a matched comparison

group design to examine the impact of the PEERS program on students' academic performance and retention in science majors during their first 2 yr at UCLA. The quasi-experimental study (Shadish *et al.*, 2002), which received human subjects approval, included two groups: 1) PEERS program participants ($n = 147$) and 2) a comparison group of students who were eligible for the PEERS program but did not to participate ($n = 533$). All students in the study entered UCLA during Fall quarter of 2009 or 2010.

The matched comparison group was generated using three key PEERS eligibility criteria: incoming science major, Scholastic Aptitude Test (SAT) math score, and a composite value of "life challenge" factors. With a scale of zero to 12, this composite measure is generated for all UCLA undergraduate applicants using institutional data on high school characteristics, parental income and education, and other relevant background variables for internal monitoring and the development of student support programs and services. The use of the life challenges composite variable allowed us to match students on a metric that encompasses a variety of socioeconomic parameters and helps to ensure that results are not driven by underlying socioeconomic disparities. In addition, this method of defining a control group cast the widest net, allowing us to compare the PEERS students with the entire UCLA population of similarly prepared students.

Student information was collected from existing campus data sources. We used admissions data to determine students' term of entry to UCLA, enrollment status, and declared major, as well as background characteristics such as gender, race/ethnicity, high school GPA, and SAT scores. We used registrar data to collect term GPA, cumulative GPA, science course enrollment, academic major, and grades in specific chemistry and math courses. Table 1 provides demographic information about both the PEERS and control

groups. To ensure the highest level of similarity between the two groups, we omitted cases with missing data for eligibility categories from the study. Also, a small number of students who entered UCLA with a life or physical sciences major but never completed a science course were removed from the control group in an effort to better match the PEERS participants.

All students included in the study intended to pursue a life or physical sciences major when they applied for admission to UCLA. A majority of PEERS participants in this study were women (68%), students from URM backgrounds (78%), and life sciences majors (68%). The mean weighted high school GPA was 4.15, and mean SAT and SAT math scores were 1797 and 623, respectively. Similarly, the control group was 64% women, 70% URM, and 66% life sciences majors. Their mean GPA was 4.11, and mean SAT and SAT math scores were 1769 and 615, respectively.

Pearson chi-square tests and one-way analysis of variance (ANOVA) were conducted to test for differences between the PEERS and control groups and between the 2009 and 2010 cohorts based on program eligibility criteria (see Table 1). There were no significant differences found between the groups at the $p < 0.05$ level for SAT math score ($F(3, 672) = 0.83, p = 0.48$) or percentage of URM students ($\chi^2(3, n = 680) = 3.35, p = 0.34$). In addition, no significant differences were found in the percentage of women ($\chi^2(3, n = 680) = 1.59, p = 0.66$), percentage of life sciences majors ($\chi^2(3, n = 680) = 0.51, p = 0.92$), and mean SAT score ($F(3, 675) = 1.78, p = 0.15$). Welch F and Brown-Forsythe tests were run to account for unequal variance, and no significant differences were found between the groups for life challenge factors ($F(3, 422.73) = 1.05, p = 0.33$) and weighted high school GPA ($F(3, 399.30) = 1.62, p = 0.15$). While this comparison group design does not fully account for self-selection of students into the PEERS program, the control group does include a number of students who

Table 1. Demographic characteristics of study participants

	Full comparison group			Propensity-matched comparison group	
	PEERS ($n = 147$)	Control ($n = 533$) ^a	High SAT math control ($n = 167$) ^b	PEERS ($n = 141$)	Matched control ($n = 141$)
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Gender (female)	0.68 (0.47)	0.64 (0.48)	0.50* (0.50)	0.68 (0.47)	0.68 (0.47)
High school GPA ^c	4.15 (0.19)	4.11 (0.22)	4.15 (0.16)	4.15 (0.19)	4.15 (0.18)
Life challenge factors ^d	4.84 (2.99)	5.29 (3.42)	4.62 (3.40)	4.82 (3.05)	5.15 (3.41)
SAT ^e	1797 (158)	1769 (168)	1903* (140)	1795 (157)	1779 (172)
SAT math ^f	623 (66)	615 (59)	684* (34)	622 (66)	622 (61)
Science division (life) ^g	0.68 (0.47)	0.66 (0.48)	0.57 (0.50)	0.69 (0.47)	0.71 (0.46)
URM ^h	0.78 (0.42)	0.70 (0.46)	0.59* (0.49)	0.79 (0.41)	0.80 (0.40)

* $p < 0.05$, difference in relation to full PEERS cohort ($n = 147$).

^aControl group includes all students who were eligible to participate in PEERS but either did not apply, were not selected, or chose not to participate.

^bHigh SAT math control group includes only those members of the full control group who earned an SAT math score of 650 or higher.

^cWeighted grade point average. Scale: 3.06–4.50.

^dScale: 0–12. A higher score means greater life challenge.

^eScale: 600–2400

^fScale: 200–800

^gIncoming science major: life or physical science

^hURM includes African-American/black, Latino/a, Chicano/a, Native American, Pacific Islander students.

applied to and were selected for PEERS but chose not to participate for various reasons or who applied to and participated in other academic support programs at UCLA.

Comparison with High Academic Achievers. To gain a better understanding of how PEERS participants performed in relation to students with higher levels of academic preparation, we identified a smaller subset of students from the control group to use as a secondary matched comparison group. Specifically, we chose those members of the control group who earned an SAT math score of 650 or higher ($n = 167$), subsequently referred to as “high SAT math control.” Because SAT score has been shown to be a strong predictor of academic success (Burton and Ramist, 2001; Cohn *et al.*, 2004) and persistence in college (Duncan and Dick, 2000; Burton and Ramist, 2001; Reason, 2009), the high SAT math comparison group can be viewed as students most prepared to succeed in STEM majors at UCLA. As expected, the high SAT math control group is significantly different from the PEERS group on various measures, with fewer females (50%) and URM students (59%) and higher mean SAT (1903) and SAT math (684) scores (see Table 1; $p < 0.05$). The results of our statistical analyses between PEERS participants and members of the high SAT math control subgroup were used to add context to our study and control for some degree of selectivity among PEERS participants, more than to assert differences between the two groups.

Propensity Score Analysis. Because a fully randomized control was not possible for this study, we used a third control group comparison employing propensity score analysis using binary logistic regression to control for the potential effects of self-selection (Rubin, 2006; Guo and Fraser, 2010). The regression model used in this study included PEERS eligibility criteria along with other relevant demographic variables: status as an URM student, high school GPA, and gender. We utilized nearest-neighbor matching with a caliper of $\epsilon \leq 0.25\sigma_p$, where σ_p denotes the SD of the estimated propensity scores in the sample (Guo and Fraser 2010), to create two matched comparison groups: matched PEERS ($n = 141$) and matched control ($n = 141$). As expected, the propensity-matched PEERS and control groups were nearly identical, with 68% in each group being female, mean SAT math score of 622, and mean weighted high school GPA of 4.15 (Table 1). Seventy-nine percent of the students from the matched PEERS group and 80% of the matched control group were URMs, and 69 and 71%, respectively, were life sciences majors when they entered UCLA.

Impact of PEERS on Academic Performance

We used Pearson chi-square tests, independent-samples t tests, and one-way ANOVA to look for significant differences between members of PEERS and the various control groups in academic performance in select lower-division chemistry and math courses supported by collaborative-learning workshops. Significance was determined using a threshold of $p < 0.05$ and by employing a Bonferroni correction to protect against type I error, resulting in a modified significance threshold of $p < 0.0046$.

Individual Course Grade Comparisons. PEERS students performed significantly ($p < 0.05$ to $p < 0.001$) better in nearly all

gatekeeper science courses than students in the control group (Table 2). With the help of collaborative-learning workshops, PEERS students earned significantly higher grades, on average, than control group members in introductory chemistry courses during Fall, Winter, and Spring quarters of their first year. For example, during Fall quarter, PEERS students earned an average grade of 2.89 in Chemistry 14A for life sciences majors and 2.80 in Chemistry 20A for physical sciences majors, compared with 2.26 and 1.92, respectively, for the control group. They also outperformed members of the control group in nearly all math courses. Even after application of a Bonferroni correction, four of 11 values remained significant.

PEERS students performed very well in comparison with members of the high SAT math control group, earning higher average grades in all but two courses, although the differences in mean grades were not statistically significant for most chemistry and math courses. However, significance disappeared after a Bonferroni correction.

Within the propensity-matched sample, PEERS students earned higher average grades in all but one of the chemistry and math courses. While grade differences could be pronounced (e.g., PEERS students earned higher average grades in chemistry courses during Fall quarter with scores of 2.87 and 2.76, compared with 2.25 and 1.85, respectively, for the control group), stringency of the propensity matching protocol resulted in sample sizes too small for meaningful statistical analyses (e.g., $n = 3$ or 7). However, multiple regression analysis for each math and chemistry course (Table 3) showed that PEERS workshops were a significant predictor of higher course grades for chemistry courses during the Fall quarter, along with a few other courses.

GPA Comparison. In addition to science course grades, PEERS students ($n = 147$) also had significantly higher overall GPAs when compared with the full control students ($n = 533$). For this analysis, we included only those PEERS students who completed 2 yr in the program and those members of the control group who persisted in science majors to Fall of their third year. The average cumulative GPA for PEERS students was 3.22 at the end of year 1 and 3.17 at the end of year 2, compared with 3.01 and 2.94, respectively, for the matched control group (Table 4). PEERS students also had a significantly higher average GPA than the high SAT math control group during year 1 ($M = 3.08$) and year 2 ($M = 3.03$). They also outperformed members of the propensity-matched control group with a mean GPA of 3.22 compared with 3.08 (Table 4) at the end of year 1, and a GPA of 3.16 compared with 3.03 at the end of year 2. All comparisons were significant at the $p < 0.05$ to $p < 0.001$ level.

Utilizing the full PEERS ($n = 147$) and control ($n = 533$) groups, we conducted multiple regression analysis to examine cumulative GPA at the end of students' second years, controlling for certain background and environmental variables (Table 5; $R^2 = 0.17$). Membership in PEERS was predictive of a higher GPA ($B = 0.43$; $p < 0.05$), while status as an URM student ($B = -0.12$) and being a physical sciences major ($B = -0.11$) were negatively associated with cumulative GPA. In line with previous research, results indicate that students' high school GPA ($B = 0.43$) and SAT math score ($B < 0.01$) were both also predictive of strong academic performance (e.g., Cohn *et al.*, 2004).

Table 2. Comparison of mean grades in first-year chemistry and math courses supported by collaborative-learning workshops: PEERS ($n = 147$) and control groups ($n = 533$)^a

	<i>n</i>	Mean	SD	<i>t</i>	<i>df</i>
Fall quarter					
Chemistry 14A					
PEERS	87 ^b	2.89	0.76	6.03***	194.26
Control	218	2.26	0.94		
High SATM control ^c	66	2.78	0.80	0.88	151
Chemistry 20A					
PEERS	34	2.80	1.12	3.97***	121
Control	89	1.92	1.09		
High SATM control	31	2.44	0.85	1.45	63
Math 1					
PEERS	16	2.87	1.12	0.10	151
Control	137	2.89	1.00		
High SATM control	10	3.04	1.01	0.39	24
Math 31A					
PEERS	22	2.86	0.55	3.23**	56.25
Control	81	2.36	0.91		
High SATM control	27	2.58	1.05	1.12	47
Math 3A					
PEERS	49	3.26	0.74	3.48**	138
Control	91	2.74	0.89		
High SATM control	27	3.15	0.93	0.59	74
Winter quarter					
Chemistry 14B					
PEERS	75	2.86	0.62	4.57***	225
Control	152	2.41	0.73		
High SATM control	54	2.60	0.72	2.14*	127
Chemistry 20B					
PEERS	22	2.94	0.60	4.17***	63
Control	43	2.12	0.81		
High SATM control	24	2.28	0.80	3.15**	44
Math 3B					
PEERS	50	3.11	0.83	1.92	130
Control	82	2.84	0.78		
High SATM control	28	3.16	0.66	0.25	76
Spring quarter					
Chemistry 14C					
PEERS	55	2.89	0.67	2.35*	150
Control	97	2.60	0.77		
High SATM control	36	2.79	0.75	0.66	89
Chemistry 30A					
PEERS	15	3.27	0.73	2.25*	35
Control	22	2.59	1.00		
High SATM control	14	3.02	0.66	0.97	27
Math 3C					
PEERS	44	2.89	0.84	2.40*	115
Control	73	2.52	0.77		
High SATM control	24	2.70	0.84	0.85	66

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.^aCourses with enrollment of < 15 students from PEERS or control group excluded: Math 31B winter quarter, Math 3A winter quarter, and Math 32A Spring quarter.^bStudents are not required to follow a specific course sequence; grade comparisons based on enrollment among study participants for each course.^cDenotes high SAT math control group.

Multiple regression analysis utilizing the same model was also conducted for the propensity-matched PEERS and control groups (Table 6; $R^2 = 0.19$, $p < 0.05$). As with the full sample, high school GPA ($B = 0.49$) and SAT math score ($B < 0.01$) were both positively correlated with the academic performance, while being a physical sciences major was negatively associated ($B = -0.10$). In contrast, science course completion was predictive of academic performance ($B = 0.01$), although only marginally, and membership in PEERS was not a significant predictor of cumulative GPA for the more restricted propensity-matched control.

Science Course Enrollment and Persistence in Science and Math Majors

To examine the ways that the PEERS program influenced students' academic performance and persistence in science majors, we measured cumulative GPA and persistence in a science major to Fall of year 3 using descriptive statistics and multiple regression analysis. First, persistence in a science major was measured by examining students' declared academic majors at the beginning of their third year at UCLA. Approximately 90% of PEERS students were retained in science majors after 2 yr ($SD = 0.31$) compared with 70% of the full control group ($SD = 0.46$, $p < 0.001$; Figure 1). Students in the control group left science and math majors at a rate three times that of PEERS students ($\chi^2(1, 680) = 22.25$, $p < 0.01$), and students in the high SAT math control group left at a rate more than twice the rate of PEERS students ($\chi^2(1, 314) = 8.78$, $p < 0.01$). Similarly, in the propensity-matched sample, 90.1% of PEERS students persisted in science majors compared with 68.8% of the control group ($\chi^2(1, 282) = 19.54$, $p < 0.001$).

Logistic regression analysis for the full sample ($n = 680$) indicated that being a female ($B = -0.79$) and having a higher GPA at the end of year 2 ($B = -0.90$) were both negatively correlated with persistence in a science major (Table 7). We speculate that this phenomenon can be explained by the fact that non-science majors often earn a higher GPA than science majors. Membership in PEERS was not a significant predictor of persistence in this model. In contrast, logistic regression analysis for the matched sample ($n = 282$) using the same model indicated that membership in PEERS was a positive predictor of persistence in a science major, with PEERS students being nearly 2.5 times as likely to persist in science majors ($B = 0.91$; Table 8). Other significant predictors of persistence were total number of science courses completed ($B = 0.42$) and SAT math score ($B = 0.01$).

Among students who persisted in science majors during their first 2 yr at UCLA, independent-samples t tests indicate that PEERS students took significantly more ($p < 0.01$) science courses (mean = 13.79, $SD = 2.76$) than students in the control group (mean = 11.15, $SD = 3.80$) and the high SAT math control group (mean = 12.25, $SD = 3.78$; Figure 2). Within the propensity-matched sample, PEERS students took an average of 13.83 ($SD = 2.78$) science courses in their first 2 yr compared with 11.37 ($SD = 3.65$) for the control group ($t(1, 178.75) = 5.45$, $p < 0.001$; Figure 3).

Multiple regression analysis for the full sample ($n = 680$), controlling for a number of background and environmental characteristics, showed that membership in PEERS predicted that students would complete a greater number of science

Table 3. Multiple regression predicting higher grade in core math and chemistry courses in the first year

	Fall quarter				Winter quarter		Spring quarter	
	Chemistry 14A	Chemistry 20A	Math 31A	Math 3A	Chemistry 14B	Math 3B	Chemistry 14C	Math 3C
	<i>Beta</i> B (SE)	<i>Beta</i> B (SE)	<i>Beta</i> B (SE)	<i>Beta</i> B (SE)	<i>Beta</i> B (SE)	<i>Beta</i> B (SE)	<i>Beta</i> B (SE)	<i>Beta</i> B (SE)
High school GPA	0.18*	0.35*	0.32*	−0.01	0.12	0.36*	0.10	0.10
	0.84 (0.33)	2.86 (0.90)	1.24 (0.48)	−0.24 (0.51)	0.37 (0.25)	1.64 (0.45)	0.34 (0.35)	0.45 (0.44)
SAT math	0.36*	0.37*	0.33*	0.17	0.39*	0.31*	0.22*	0.52*
	0.01 (<0.01)	0.01 (<0.01)	<0.01 (<0.01)	<0.01 (<0.01)	<0.01 (<0.01)	<0.01 (<0.01)	<0.01 (<0.01)	0.01 (<0.01)
Gender (female)	−0.04	0.03	0.02	−0.24*	0.04	0.04	−0.08	−0.06
	−0.08 (0.15)	0.08 (0.26)	0.03 (0.17)	−0.44 (0.20)	0.05 (0.11)	0.07 (0.19)	−0.12 (0.16)	−0.13 (0.22)
URM	−0.11	−0.09	−0.24	−0.10	−0.04	0.01	<0.01	0.04
	−0.27 (0.17)	−0.25 (0.30)	−0.35 (0.19)	−0.19 (0.23)	−0.05 (0.31)	0.02 (0.22)	<0.01 (0.17)	0.07 (0.22)
Incoming science division (physical)	0.04	−0.14	−0.21	0.16	−0.01	0.01	−0.05	0.07
	0.11 (0.22)	−0.43 (0.32)	−0.27 (0.16)	0.31 (0.21)	−0.02 (0.16)	0.03 (0.29)	−0.10 (0.23)	0.19 (0.28)
Participation in PEERS Workshop	0.31*	0.28*	0.19	0.24*	0.22*	0.15	0.04	0.14
	0.57 (0.13)	0.69 (0.26)	0.24 (0.16)	0.39 (0.18)	0.27 (0.10)	0.24 (0.16)	0.06 (0.14)	0.25 (0.18)
R ²	0.30	0.48	0.46	0.18	0.21	0.24	0.07	0.27

* $p < 0.05$.

courses during the first 2 yr of college ($B = 2.85$; Table 9). In addition, SAT math score ($B = 0.02$) and cumulative GPA ($B = 2.94$) were both positively correlated with the dependent variable, while status as a URM student was negatively correlated ($B = -0.85$). The full model accounted for ~20% of the variance in science course enrollments ($R^2 = 0.20$).

Utilizing the same regression model for the matched PEERS and control groups ($n = 282$), results indicate that membership in PEERS was a predictor of higher science course completion ($B = 2.86$; Table 10). Cumulative GPA at the end of year 2 ($B = 1.32$) and SAT math score ($B = 0.02$) were also significant predictors of completing more science courses. The full model accounted for nearly 30% of the

variance in course completion rates ($R^2 = 0.29$), indicating that even among a population of similar students, membership in PEERS provided an academic and social support system that encouraged timely progress toward the successful completion of an undergraduate science degree.

DISCUSSION

The persistence of URM students in STEM majors is a function of their confidence and motivation (Graham *et al.*, 2013). However, the confidence and motivation of many STEM majors from underrepresented groups are challenged

Table 4. Cumulative GPA comparison of PEERS to nonmember control groups^a

	<i>n</i>	Mean GPA	SD	<i>t</i>	<i>df</i>
Cumulative GPA Spring year 1					
PEERS	112	3.22	0.41	4.75***	485
Control	375	3.01	0.43		
High SAT math control	129	3.08	0.44	2.60**	239
Cumulative GPA Spring year 2					
PEERS	112	3.17	0.43	5.02***	478
Control	368	2.94	0.43		
High SAT math control	128	3.03	0.47	2.37*	238
Cumulative GPA Spring year 1					
Matched PEERS	107	3.22	0.45	2.29*	201
Matched control	96	3.08	0.40		
Cumulative GPA Spring year 2					
Matched PEERS	107	3.16	0.40	2.39*	200
Matched control	95	3.03	0.40		

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

^aIncludes only PEERS students who completed 2 yr in the program and members of the control groups who persisted in science majors to Fall year 3.

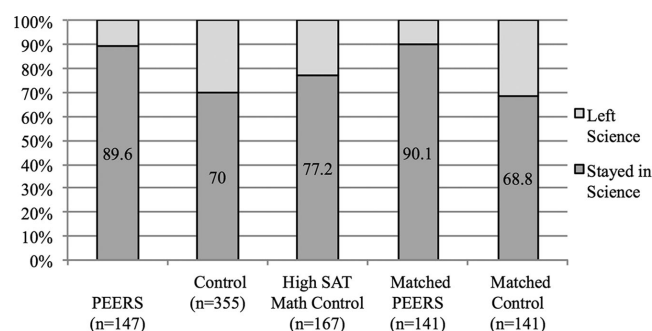
Table 5. Multiple regression predicting cumulative GPA at the end of year 2 for PEERS ($n = 147$) and control ($n = 533$) groups

	Standardized coefficients	Unstandardized coefficients	
	Beta	B	SE B
Constant		0.43	0.36
High school GPA (pretest)	0.21	0.43*	0.08
Gender (female)	0.02	0.01	0.03
URM	-0.13	-0.12*	0.04
SAT math	0.20	<0.01*	0.00
Membership in PEERS	0.12	0.13*	0.04
Incoming science division (physical)	-0.12	-0.11*	0.03
Total number of science courses	0.05	0.01	0.00
$R^2 = 0.166$			
* $p < 0.05$.			

as they transition to college (Cooper *et al.*, 2005) and encounter an academic culture perceived as unwelcoming (Ong *et al.*, 2011; Beasley and Fischer, 2012). Becoming an active member of a learning community (Tinto, 1997; Museus, 2011; Walton and Cohen, 2011), experiencing positive social and academic integration (Tinto, 1987; Astin, 1993; Museus and Quaye, 2009), and engaging in undergraduate research (Lopatto, 2007; Carter *et al.*, 2009) can each play an important role in URM students' persistence and ability to envision themselves as scientists. Conversely, students who struggle to complete introductory science and math courses may lose confidence in their abilities and motivation to persist, resulting in attrition from STEM majors (Labov, 2004; Gasiewski *et al.*, 2012). Thus, it follows that an institutional commitment to promoting an inclusive campus culture and offering focused academic support programs can increase persistence of students from underrepresented backgrounds (Matsui *et al.*, 2003; Muesus, 2011; Stolle-McAllister *et al.*, 2011).

Table 6. Multiple regression predicting cumulative GPA at the end of year 2 for matched PEERS ($n = 141$) and control ($n = 141$) groups

	Standardized coefficients	Unstandardized coefficients	
	Beta	B	SE B
Constant		0.21	0.61
High school GPA (pretest)	0.22	0.49*	0.13
Gender (female)	0.07	0.06	0.05
URM	-0.10	-0.10	0.06
SAT math	0.19	<0.01*	<0.01
Membership in PEERS	0.08	0.06	0.05
Incoming science division (physical)	-0.11	-0.10*	0.05
Total number of science courses	0.15	0.01*	0.01
$R^2 = 0.186$			
* $p < 0.05$.			

**Figure 1.** The percentage of students in each group—PEERS, control group, high SAT math control, propensity-matched PEERS, and matched control—who were retained in science majors as of the Fall quarter of their third undergraduate year. Science majors include life sciences, physical sciences, and engineering.

PEERS provides targeted support to students who come from groups typically associated with lower levels of persistence in science majors, including women, URMs, and students from lower socioeconomic backgrounds (Koenig, 2009; Hurtado *et al.*, 2010; PCAST, 2012). Yet results indicate that PEERS provides the necessary support system for participants to achieve higher levels of academic success and persistence. Longitudinal comparison of students in the UCLA PEERS program shows that being a member of PEERS had a significant positive impact on multiple measures of academic success.

With the help of Treisman-style collaborative-learning workshops (Treisman, 1992), PEERS students, on average, earned higher grades than non-PEERS students in the foundational introductory math and science courses that form the core science curriculum at UCLA—the gatekeeper courses that frequently pose an academic barrier to incoming science students (Fries-Britt *et al.*, 2010; Gasiewski *et al.*, 2012). Supplemental instruction of this kind in introductory science courses has been shown to increase persistence and

Table 7. Logistic regression predicting retention in science major to Fall year 3 for PEERS ($n = 147$) and control ($n = 533$) groups

	Unstandardized coefficients		Odds ratio
	B	SE B	Exp(B)
Constant	-3.97	2.42	
Gender (female)	-0.79*	0.27	0.46
URM	0.51	0.27	1.66
High school GPA	0.44	0.53	1.55
SAT math	0.00	0.00	1.00
Membership in PEERS	0.63	0.37	1.89
Incoming science division (physical)	0.19	0.25	1.21
Cumulative GPA Spring year 2	-0.90*	0.28	0.41
Total science courses completed during first 2 yr	0.41*	0.04	1.51
$R^2 = 0.332$			
* $p < 0.05$.			

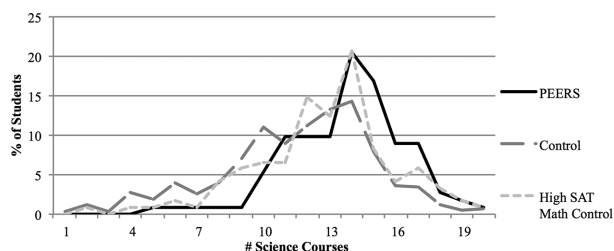
Table 8. Logistic regression predicting retention in science major to Fall year 3 for matched PEERS ($n = 141$) and control ($n = 141$) groups

	Unstandardized coefficients		Odds ratio
	B	SE B	Exp(B)
Constant	-11.90	5.24	
Gender (female)	-0.15	0.50	0.86
URM	0.32	0.56	1.38
High school GPA	0.76	1.11	2.14
SAT math	0.01*	0.00	1.01
Membership in PEERS	0.91*	0.45	2.49
Incoming science division (physical)	0.87	0.47	2.38
Cumulative GPA Spring year 2	-0.61	0.51	0.54
Total science courses completed during first 2 yr	0.42*	0.07	1.52

 $R^2 = 0.348$ * $p < 0.05$.

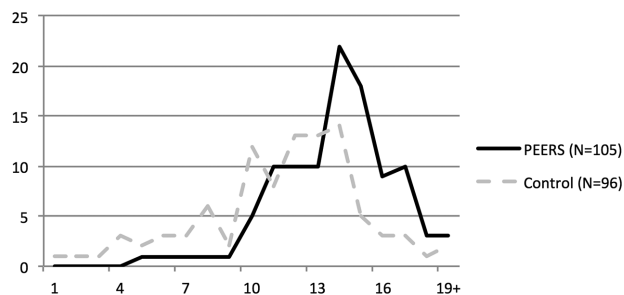
academic success of women and URM students (Peterfreund *et al.*, 2007). In addition, these workshops promote more active learning, which contributes to improved academic engagement (Handelsman *et al.*, 2005; Derting and Ebert-May, 2010).

For PEERS students, improved performance in their science courses translated to higher first- and second-year GPAs and higher enrollments on average in science courses during the first 2 yr, as compared with members of the three different control groups. In addition, 90% of PEERS students persisted in science majors during their first 2 yr, a rate significantly higher than those of the control groups. These findings suggest that participation in the PEERS program provided students with



Note: Includes only PEERS students who completed the program and Control students who stayed in science majors through Fall Year 3.

Figure 2. The total number of science courses that students completed during their first 2 yr at UCLA. Course number is based on tally of student enrollment in life and physical sciences, mathematics, and statistics courses during each quarter of their first and second year. Lines depict the percentage of students in each group—PEERS, control, and high SAT math control—who completed a specific number of science courses. PEERS students took an average of 13.79 science courses ($SD = 2.76$), while students in the control group took an average of 11.15 ($SD = 3.80$). On average, students in the high SAT math control subgroup enrolled in 12.25 science courses ($SD = 3.78$). Independent-samples t tests indicate that PEERS students took significantly more science courses during their first 2 yr than members of the control group ($t(1, 247.97) = 8.07, p < 0.001$), and the high SAT math control group ($t(1, 232.56) = 3.64, p < 0.001$).



Note: Includes only PEERS and Control students who stayed in science majors through Fall Year 3.

Figure 3. The total number of science courses that students completed during their first 2 yr at UCLA as described in Figure 2. Lines depict the percentage of students in each propensity-matched group—matched PEERS and matched control. PEERS students took an average of 13.83 ($SD = 2.78$) science courses in their first 2 yr compared with 11.37 ($SD = 3.65$) for the control group ($t(1, 178.75) = 5.45, p < 0.001$).

the tools to succeed academically, which in turn may have encouraged higher levels of confidence and motivation to pursue a STEM degree among participants. Comparing PEERS participants with the high SAT math control group provides additional insights regarding the impact of the program on academic performance and persistence. The high SAT math control group represents those members of the control group with the best academic preparation, as measured by SAT math score. Yet PEERS students outperformed this group on numerous measures. Not only did PEERS students earn grades in their math courses that were comparable to those of well-prepared high SAT math control students, PEERS students outperformed this group in a number of chemistry courses, had significantly higher cumulative GPAs, took more science courses, and left STEM majors at a rate that was 50% lower. Thus, the PEERS program helped to close the achievement gap by supporting students' academic performance and persistence in STEM majors despite the numerous risk factors typically associated with higher levels of attrition.

Table 9. Multiple regression predicting total number of science courses completed at the end of year 2 for PEERS ($n = 147$) and control ($n = 533$) groups

	Standardized coefficients	Unstandardized coefficients	
	Beta	B	SE B
Constant		-14.75	3.47
High school GPA (pretest)	0.14	2.94*	0.75
Gender (female)	-0.03	-0.26	0.33
URM	-0.09	-0.85*	0.35
SAT math	0.24	0.02*	0.00
Membership in PEERS	0.27	2.85*	0.38
Incoming science division (physical)	-0.03	-0.24	0.33
Total number of science courses	0.05	0.49	0.38

 $R^2 = 0.200$ * $p < 0.05$.

Table 10. Multiple regression predicting total number of science courses completed at the end of year 2 for matched PEERS ($n = 141$) and control ($n = 141$) groups

	Standardized coefficients	Unstandardized coefficients	
	Beta	B	SE B
Constant		-11.99	5.73
High school GPA (pretest)	0.05	1.08	1.23
Gender (female)	-0.08	-0.75	0.48
URM	-0.03	-0.30	0.56
SAT math	0.30	0.02*	0.00
Membership in PEERS	0.34	2.86*	0.44
Incoming science division (physical)	-0.01	-0.10	0.48
Cumulative GPA Spring year 2	0.13	1.32*	0.57

$R^2 = 0.292$

* $p < 0.05$.

Much assessment related to STEM support programs relies on students' self-reported outcomes or provides information about the members' achievements in comparison with institutional or national averages. While these are both useful measures, the application of existing institutional data sources provided us with a unique opportunity to create a matched comparison group and consider the impact of PEERS on students' academic performance and persistence in science and math majors. Results demonstrate that membership in PEERS confers significant benefits to students that contribute to increased persistence. However, our study design does not allow us to determine which components of the program convey the greatest positive impact. PEERS students participate in a seminar course focused on developing study and time-management skills, and they receive specialized counseling focused on creating a yearly schedule of classes to match their individual backgrounds and academic goals. Both of these activities could contribute to the higher course grades and cumulative GPA of PEERS students in comparison with the control groups. Similarly, holistic counseling and access to collaborative-learning workshops could encourage students to take more science courses in their first 2 yr (while they have access to these workshops), and the sense of community developed in PEERS likely contributes to improved retention (Museus and Quayle, 2009; Walton and Cohen, 2011). Finally, the seminar focused on career preparation could help PEERS students envision themselves in a scientific career, promoting increased persistence in science majors.

PEERS students outperformed the full control group and high SAT math control group in every comparison, ranging from individual courses to cumulative GPA, number of science courses, and persistence in science majors. However, our study design is potentially susceptible to the effects of self-selection, wherein higher academic performance results from the motivation bias in students who apply to a program, rather than from the beneficial effects of program activities (Kohn *et al.*, 1976; Heckman, 1979; Eagan *et al.*, 2013). However, multiple lines of evidence indicate that this is not the case. First, due to funding limitations, more than

200 applicants are turned away from PEERS each year. Thus, the control group ($n = 533$) includes a substantial number of students who were motivated to apply to PEERS but were not admitted. Second, many students in the control group are part of other academic support programs at UCLA that also require students to apply. Finally, even when we used a strict propensity score analysis designed to control for some degree of self-selection bias (Rubin, 2006; Guo and Fraser, 2010), results showed that PEERS students still displayed stronger academic performance, higher grades in many science classes, and substantially lower rates of attrition from science majors. Together, these factors indicate that every effort was made to control for self-selection bias and examine the influence of the PEERS program on students' academic performance and persistence.

While direct comparisons of academic performance between the groups indicated that PEERS students earned higher average cumulative GPAs, the results of the regression models are more nuanced. Using the full control group, membership in PEERS was a significant predictor of a higher cumulative GPA and taking more science classes. However, PEERS membership was not predictive of persistence in a science major, even though PEERS students take significantly more science courses and the regression model indicates that total number of science courses is predictive of persistence. Similarly, the propensity-matched comparison showed that membership in PEERS was a significant predictor of total number of science courses and persistence in science majors but was not a predictor of higher cumulative GPA, even though PEERS students have significantly higher GPAs at the end of years 1 and 2.

One potential explanation of the variable results in the regression models is attrition from science majors within the control group. Using the Chemistry 14 series as an example, the control group included 218 students in the first quarter, 152 in the second, and 97 in the third, a reduction of 30.3 and 55.5%, respectively. The number of enrolled students in the control group shrinks, because students in the control group leave STEM majors, which is likely a result of poor performance in their core science courses. As such, mean performance of the control group is higher than it would be if students in the lower tail of the distribution had remained in STEM.

Similarly, there is a decrease in the number of PEERS students taking the collaborative-learning workshops (e.g., 87, 75, and 55 students in the first, second, and third quarters, respectively, of the Chemistry 14 series). However, this does not change the number of students in the PEERS treatment group. PEERS students enrolled in the full three quarters of the Chemistry 14 series but, in some cases, did not enroll in the accompanying collaborative-learning workshops. PEERS students are not required to take every workshop associated with their core science curriculum. Thus, while fewer students take the collaborative-learning workshops as they become more confident with the material or prioritize other workshop offerings in math or physics, such decreases do not reflect a real decrease in PEERS membership. It is difficult to quantify how quarter-by-quarter changes in membership of the control group combined with stability of the PEERS treatment group impacts the regression models, but it almost certainly has an impact. Despite this issue, the models still indicate a positive impact of PEERS membership in the majority of comparisons.

Toward Best Practices

A major goal of this paper is to provide a framework to guide assessment of academic support programs based on student performance data. While the use of both the Bonferroni correction and propensity matching adds additional statistical rigor (and also controls for type I error and self-selection, respectively), the results of these analyses also highlight how difficult it is to evaluate support programs. The Bonferroni correction of course grade comparison data resulted in the loss of statistical significance in many comparisons; four of 11 course comparisons that remained significant either represented comparisons with the largest sample sizes (e.g., $n > 75$) or the biggest grade differences (mean GPA difference > 0.8). Thus, while it is important to control for type I error, a Bonferroni correction can result in significant programmatic impacts being dismissed as statistical artifacts, resulting in a type II error, particularly when differences are modest and/or sample sizes are low (Perneger, 1998; Nakagawa, 2004). As such, this test may not be appropriate for analyses for academic support programs, particularly those with smaller enrollments.

Similarly, propensity score matching in course-by-course comparisons resulted in sample sizes that were too small for meaningful statistical comparison. None of the propensity-matched grade comparisons was significant, even though PEERS students outperformed the propensity-matched control. When using the full propensity-matched control group for the GPA comparisons, however, higher sample size ($n = 147$) results in significant values, showing that PEERS students outperform the control. Similarly, propensity score-matching comparisons with higher sample sizes show that PEERS students took significantly more science courses. Combined, these results indicate that positive impacts of PEERS reported are not being driven by self-selection. Thus, while propensity score matching can be valuable in certain circumstances, the stringency of this analysis also creates significant limitations. Specifically, creating perfectly matched experimental and control groups from students representing diverse backgrounds and educational objectives can result in statistically meaningless sample sizes, suggesting that propensity score matching may have limited utility in all but the largest academic support programs.

The results of this study argue that, while statistical, data-based assessment of academic support programs is essential to demonstrate programmatic impacts, care and judgment must be applied in their application and interpretation. In our case, the reduction in significant comparisons across the math and science courses following a Bonferroni correction should not be viewed as PEERS having an impact in a smaller number of courses—while one or a few significance values could potentially be explained away by increased type I error in multiple comparisons, they cannot all be explained away. Indeed, across the board, PEERS students performed better than the control in 10 of 11 comparisons, a result that would be observed at random at a level of $p < 0.001$. Given the fact that we observed similar results using the larger pool of eligible students as a control group, the high math SAT control group and the propensity-matched controls, our findings are clearly not an artifact of multiple comparison. Similarly, results of course-by-course grade comparisons using propensity score-matched controls should not be interpreted as our results being driven by self-selection. Instead,

this result is clearly driven by lack of statistical power resulting from low sample sizes—an artifact of the propensity score-matching protocol. Combined, these results show that a variety of approaches may need to be used when evaluating a complex academic support program and that care must be applied in interpreting results so that significant programmatic impacts are not overlooked.

While this study highlights the impact of the PEERS program on students' academic success and persistence in science majors during their first 2 yr at UCLA, further study is needed to more fully understand the impact of the program on graduation rates and entry into graduate programs in science. Previous studies show that URM students who participate in targeted academic support programs in the sciences are retained in science majors to graduation at rates similar to their non-URM peers (Matsui *et al.*, 2003), and pursue graduate education at greater rates (Maton *et al.*, 2009). Since the PEERS program was established in 2003, 83% of PEERS students who completed their degrees at UCLA graduated with a science major within 5 yr. Compared with a persistence rate of 39% for URM graduates in science majors at UCLA, and a national average of 20% in 2009 (Hurtado *et al.*, 2010), the PEERS program has encouraged much higher graduation rates among participants—an outcome that deserves further exploration.

CONCLUSIONS

A variety of social, economic, and educational factors contribute to the large and stubborn gap in persistence of URM and non-URM students in science majors. Research has shown that a diversity of interventions that foster supportive peer networks, create a more welcoming academic culture, and allow students to begin to see themselves as scientists can all help improve retention of underrepresented students. However, it is also imperative to increase the academic confidence and motivation of these students (Graham *et al.*, 2013). Through a holistic approach that focuses on academics, counseling, creating a supportive community, and exposure to research, the PEERS program has a strong positive effect, resulting in academic performance and persistence that exceeds even the most academically prepared students from similar backgrounds. As such, PEERS serves as an excellent model for universities interested in and committed to improving persistence of underrepresented science majors and closing the achievement gap.

ACKNOWLEDGMENTS

The PEERS program is indebted to C. Buchanan and R. Weiss for their vision and leadership and to the PEERS advisory board members for their guidance. We thank Dean Judi Smith, Dean Pat Turner, and the Division of Undergraduate Education at UCLA for their ongoing support of the PEERS Program. Research reported in this publication was supported by an award from the National Institute of General Medical Sciences of the National Institutes of Health under award number R25GM055052, an NSF Louis Stokes Alliances for Minority Participation award to the California Alliance for Minority Participation (NSF1102531 to Drake—University of California, Irvine), a Howard Hughes Medical Institute Undergraduate Science Education Grant (Eiserling, P.I.), and the Division of Undergraduate Education at UCLA.

REFERENCES

- Astin AW (1993). *What Matters in College? Four Critical Years Revisited*, San Francisco, CA: Jossey-Bass.
- Barefoot BO (2000). The first-year experience. *About Campus* 4, 612–18.
- Barlow AE, Villarejo M (2004). Making a difference for minorities: evaluation of an educational enrichment program. *J Res Sci Teach* 41, 861–881.
- Beasley MA, Fischer MJ (2012). Why they leave: the impact of stereotype threat on the attrition of women and minorities from science, math and engineering majors. *Soc Psychol Educ* 15, 427–448.
- Bonous-Hammarth M (2000). Pathways to success: affirming opportunities for science, mathematics, and engineering majors. *J Negro Educ* 69, 92–111.
- Burton NW, Ramist L (2001). *Predicting Success in College: SAT Studies of Classes Graduating since 1980*, New York: College Entrance Examination Board.
- Carter FD, Mandell M, Maton KI (2009). The influence of on-campus, academic year undergraduate research on STEM Ph.D. outcomes: evidence from the Meyerhoff scholarship program. *Educ Eval Policy Anal* 31, 441–462.
- Chang MJ, Cerna O, Han J, Saenz V (2008). The contradictory roles of institutional status in retaining underrepresented minorities in biomedical and behavioral science majors. *Rev High Educ* 31, 433–464.
- Chang MJ, Eagan MK, Lin MH, Hurtado S (2011). Considering the impact of racial stigmas and science identity: persistence among biomedical and behavioral science aspirants. *J High Educ* 82, 564.
- Chang MJ, Sharkness J, Hurtado S, Newman CB (2014). What matters in college for retaining aspiring scientists and engineers from underrepresented racial groups. *J Res Sci Teach* 51, 555–580.
- Choy SP, Horn LJ, Nuñez AM, Chen X (2000). Transition to college: what helps at-risk students and students whose parents did not attend college. *New Dir Inst Res* 2000, 45–63.
- Cohn E, Cohn S, Balch DC, Bradley J, Jr (2004). Determinants of undergraduate GPAs: SAT scores, high-school GPA and high-school rank. *Econ Educ Rev* 23, 577–586.
- Cooper CR, Chavira G, Mena DD (2005). From pipelines to partnerships: a synthesis of research on how diverse families, schools, and communities support children's pathways through school. *J Educ Stud Placed Risk* 10, 407–430.
- Crisp G, Nora A, Taggart A (2009). Student characteristics, pre-college, college, and environmental factors as predictors of majoring in and earning a STEM degree: an analysis of students attending a Hispanic serving institution. *Am Educ Res J* 46, 924–942.
- DeAngelo L, Hasson T (2009). Quantifying success: using control groups to measure program effectiveness. *Counc Undergrad Res* 29, 39–45.
- Derting TL, Ebert-May D (2010). Learner-centered inquiry in undergraduate biology: positive relationships with long-term student achievement. *CBE Life Sci Educ* 9, 462–472.
- Duncan H, Dick T (2000). Collaborative workshops and student academic performance in introductory college mathematics courses: a study of a Treisman model math excel program. *School Sci Math* 100, 365–373.
- Eagan MK Jr, Hurtado S, Chang MJ, Garcia GA, Herrera FA, Garibay JC (2013). Making a difference in science education: the impact of undergraduate research programs. *Am Educ Res J* 50, 683–713.
- Elliott R, Strenta AC, Adair R, Matier M, Scott J (1996). The role of ethnicity in choosing and leaving science in highly selective institutions. *Res High Educ* 37, 681–709.
- Fries-Britt SL, Younger TK, Hall WD (2010). Lessons from high-achieving students of color in physics. *New Dir Inst Res* 148, 75–83.
- Gasiewski JA, Eagan MK, Garcia GA, Hurtado S, Chang MJ (2012). From gatekeeping to engagement: a multi-contextual, mixed method study of student academic engagement in introductory STEM courses. *Res High Educ* 53, 229–261.
- Graham MJ, Frederick J, Byars-Winston A, Hunter A, Handelsman J (2013). Increasing persistence of college students in STEM. *Science* 341, 1455–1456.
- Guo S, Fraser MW (2010). *Propensity Score Analysis: Statistical Methods and Applications*, vol. 11, Thousand Oaks, CA: Sage Publications.
- Handelsman MM, Briggs WL, Sullivan N, Towler A (2005). A measure of college student course engagement. *J Educ Res* 98, 184–192.
- Heckman JJ (1979). Sample selection bias as a specification error. *Econometric Soc* 47, 153–161.
- Hira R (2010). US Policy and the STEM workforce system. *Am Behav Sci* 53, 949–961.
- Hurtado S, Eagan K, Chang M (2010). *Degrees of Success: Bachelor's Degree Completion Rates among Initial STEM Majors*, Los Angeles, CA: Higher Education Research Institute at UCLA.
- Johnson AC (2007). Unintended consequences: how science professors discourage women of color. *Sci Educ* 91, 805–821.
- Koenig R (2009). Minority retention rates in science are sore spot for most universities. *Science* 324, 1386–1387.
- Kohn MG, Mansk CF, Mundel DS (1976). An empirical investigation of factors which influence college-going behavior. *Ann Econ Soc Meas* 5, 391–419.
- Labov JB (2004). The challenges and opportunities for improving undergraduate science education through introductory courses. *Cell Biol Educ* 3, 212–214.
- Lopatto D (2007). Undergraduate research experiences support science career decisions and active learning. *CBE Life Sci Educ* 6, 297–306.
- Maton KI, Domingo MRS, Stolle-McAllister KE, Zimmerman JL, Hrabowski FA III (2009). Enhancing the number of African-Americans who pursue STEM PhDs: Meyerhoff Scholarship Program outcomes, processes, and individual predictors. *J Women Minorities Sci Eng* 15, 15–37.
- Matsui J, Liu R, Kane CM (2003). Evaluating a science diversity program at UC Berkeley: more questions than answers. *Cell Biol Educ* 2, 117–121.
- McCarron GP, Inkelas KK (2006). The gap between educational aspirations and attainment for first-generation college students and the role of parental involvement. *J Coll Stud Dev* 47, 534–549.
- Museus SD (2011). Generating Ethnic Minority Student Success (GEMS): a qualitative analysis of high-performing institutions. *J Divers Higher Educ* 4, 147.
- Museus SD, Quaye SJ (2009). Toward an intercultural perspective of racial and ethnic minority college student persistence. *Rev High Educ* 33, 67–94.
- Nakagawa S (2004). A farewell to Bonferroni: the problems of low statistical power and publication bias. *Behav Ecol* 15, 1044–1045.
- National Academy of Sciences, National Academy of Engineering, and Institute of Medicine (2011). *Expanding Underrepresented Minority Participation: America's Science and Technology Talent at the Crossroads*, Washington, DC: National Academies Press.
- National Board of Sciences (2010). *Preparing the Next Generation of STEM Innovators: Identifying and Supporting Our Nation's Human Capital*, Arlington, VA: NSF Publications. www.nsf.gov/nsb/publications/2010/nsb1033.pdf (accessed 29 December 2012).
- National Science Foundation (2012). *Women, Minorities, and Persons with Disabilities in Science and Engineering: Data Tables*. www.nsf.gov/statistics/wmpd/tables.cfm (accessed 29 December 2012).

- Ong M, Wright C, Espinosa LL, Orfield G (2011). Inside the double bind: a synthesis of empirical research on undergraduate and graduate women of color in science, technology, engineering, and mathematics. *Harvard Educ Rev* 81, 172–209.
- Perneger TV (1998). What's wrong with Bonferroni adjustments. *BMJ* 316, 1236–1238.
- Peterfreund AR, Rath KA, Xenos SP, Bayliss F (2007). The impact of supplemental instruction on students in STEM courses: results from San Francisco State University. *J Coll Stud Retent* 9, 487–503.
- President's Council of Advisors on Science and Technology (2012). *Engage to Excel: Producing One Million Additional College Graduates with Degrees in Science, Technology, Engineering and Mathematics*, Washington, DC: U.S. Government Office of Science and Technology.
- Reason RD (2009). An examination of persistence research through the lens of a comprehensive conceptual framework. *J Coll Stud Dev* 50, 659–682.
- Reason RD, Terenzini PT, Domingo RJ (2006). First things first: developing academic competence in the first year of college. *Res High Educ* 47, 149–175.
- Rubin DB (2006). *Matched Sampling for Causal Effects*, New York: Cambridge University Press.
- Shadish W, Cook T, Campbell D (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*, Belmont, CA: Wadworth Cengage Learning.
- Stolle-McAllister K, Domingo MRS, Carrillo A (2011). The Meyerhoff way: how the Meyerhoff scholarship program helps black students succeed in the sciences. *J Sci Educ Technol* 20, 5–16.
- Summers MF, Hrabowski FA III (2006). Preparing minority scientists and engineers. *Science* 311, 1870–1871.
- Terenzini PT, Springer L, Yaeger PM, Pascarella ET, Nora A (1996). First-generation college students: characteristics, experiences, and cognitive development. *Res High Educ* 37, 1–22.
- Tinto V (1987). *Leaving College: Rethinking the Causes and Cures of Student Attrition*, Chicago, IL: University of Chicago Press.
- Tinto V (1997). Classrooms as communities. *J High Educ* 68, 599–623.
- Tinto V (2005). Taking student success seriously: rethinking the first year of college. In: *Ninth Annual Intersession Academic Affairs Forum*, Fullerton: California State University, 05–01.
- Treisman PU (1992). Studying students studying calculus: a look at the lives of minority mathematics students in college. *Coll Math J* 23, 362–372.
- Walton GM, Cohen GL (2011). A brief social-belonging intervention improves academic and health outcomes of minority students. *Science* 331, 1447–1451.
- White JL, Altschuld JW, Lee YF (2008). Evaluating minority retention programs: problems encountered and lessons learned from the Ohio science and engineering alliance. *Eval Program Plann* 31, 277–283.