Improving Academic Performance and Retention of First-Year Biology Students through a Scalable Peer Mentorship Program

Mike Wilton,[†]* Daniel Katz,[‡] Anthony Clairmont,[‡] Eduardo Gonzalez-Nino[†] Kathy R. Foltz,[†] and Rolf E. Christoffersen[†]

[†]Department of Molecular, Cellular, and Developmental Biology and [‡]Gevirtz Graduate School of Education, University of California, Santa Barbara, CA 93106

ABSTRACT

We examine the impact of <u>Bio</u>logy <u>M</u>entoring and <u>E</u>ngagement (BIOME) near-peer mentorship on 437 first-year undergraduate students over three cohort years. The BIOME course consists of ten, 50-minute meetings where groups of six first-year mentees meet with an upper-division student mentor to discuss topics including metacognition, growth mindset, and effective study strategies. We employed a mixed-methods approach to evaluate the impact of BIOME on mentee academic outcomes. Initial ethnographic analysis revealed that BIOME influenced student study methods, approaches to academic challenges, and use of campus learning communities. We then constructed a novel, program-specific instrument to measure the implementation of these habits, a construct we named "academic habit complexity." Regression analysis supported the hypothesis that enrollment in BIOME leads to students using more diverse approaches than their peers. Enrollment in BIOME, and the associated development of academic habit complexity, is related to higher course grades in General Chemistry, a biology major prerequisite. Finally, students participating in BIOME demonstrated improved short-term student retention as measured by increased enrollment in the subsequent prerequisite General Chemistry course. These results suggest that course-based near-peer mentorship may be an effective and scalable approach that can promote student academic success.

INTRODUCTION

An increasing number of reports have called for changes to existing pedagogical practices and educational structures to increase the quality, number, and diversity of science, technology, engineering, and mathematics (STEM) graduates (National Research Council, 2009; 2015; President's Council of Advisors on Science and Technology [PCAST], 2012; National Academies of Sciences, Engineering, and Medicine [NASEM], 2018). This can be accomplished, in part, by retaining more individuals who enter college intending to pursue a STEM degree. The highest attrition among these students occurs early in the college experience (Tinto, 1988, 2006; Almatrafi *et al.*, 2017). Further, this attrition is disproportionately higher among students who are first-generation (first in their families to attend college) or are people of color (including Latinx, Black, Indigenous or Native American, and Pacific Islander). Loss of persons excluded because of their ethnicity or race (PEERs; historically called underrepresented minorities or URMs; Asai, 2020) leads to a loss of diversity in STEM fields that must be addressed (Griffith, 2010; Chang *et al.*, 2014; Eagan *et al.*, 2017; Riegle-Crumb *et al.*, 2019; Asai, 2020).

Students' rationales for leaving STEM programs are multifaceted, reliant on the intersection of academic experience, performance, and noncognitive perception (Hall and Sandler, 1982; Chang *et al.*, 2014; Aryee, 2017; Eagan *et al.*, 2017). Reasons include the "chilly climate" of STEM classrooms and departments that makes it difficult

to build sense of belonging, the fast-paced and content-heavy nature of the introductory curriculum, and a competitive culture that prioritizes independence (Hall and Sandler, 1982; Seymour and Hewitt, 1997; PCAST, 2012; Fabert, 2014). Not surprisingly, research has identified that retention in STEM relies on student social integration, fostered interest, and self-efficacy (Tinto, 1993; Chemers *et al.*, 2001; Hoffman *et al.*, 2002; Estrada *et al.*, 2011; Solanki *et al.*, 2019). However, given the public university context, which often includes noninteractive, large-enrollment courses, it is difficult to provide first-year students with intimate experiences that inculcate student experiences and traits that lead to improved retention (Stains *et al.*, 2018).

Theoretical Framework of Mentorship
Research exploring the relationship between undergraduate mentoring and student development is guided by a number of different theoretical frameworks (as reviewed in Crisp and Cruz, 2009; Gershenfeld, 2014). The theory most often applied to undergraduate mentorship programs in higher education is Tinto's integration framework (Tinto, 1993). Tinto posits that students who are integrated into the campus environment, both within and outside the classroom, are more likely to persist. Feelings of integration often contribute to greater satisfaction with and commitment to the university, both of which influence student retention decisions. A response to Tinto's work has been to implement structured student support services, including mentorship opportunities, meant to encourage integration.

The extant framework of mentorship suggests that this practice facilitates mentee integration in a number of ways (Jacobi, 1991; Austin, 2002; Allen and Eby, 2003; Ostrove and Long, 2007; NASEM, 2019), including psychosocial support, instrumental support, and academic support (Kram, 1988; Scandura, 1992; Nora and Crisp, 2007; Eby *et al.*, 2013). Psychosocial support refers to mentor behaviors that enhance a mentee's self-perception of competence and facilitates both personal and emotional development (Kram, 1985; Flaxman *et al.*, 1988; Tenenbaum *et al.*, 2001; Nakkula and Harris, 2005; Johnson *et al.*, 2007; Spencer, 2007). Instrumental support refers to resources and opportunities provided by the mentor to mentees that enable their engagement in achieving goals and belonging (Tenenbaum *et al.*, 2001; Allen and Eby, 2007; Blinn-Pike, 2007; Johnson *et al.*, 2007; Griffin and Romm, 2008; Terrion, 2012). Academic support involves educating, evaluating, and challenging students academically and includes subject-knowledge support, goal setting, skill development, and career advice (Schockett and Haring-Hidore, 1985; Miller, 2002; Nora and Crisp, 2007). Finally, scholars examining mentorship for youth and in the workplace have attributed high-quality connections to effective mentoring relationships (Rhodes, 2002, 2005). Thus, by combining these elements, whereby mentors make explicit effective disciplinary behaviors and practices, the model suggests that mentorship can promote the adoption of new habits that improve mentee success in specialized, challenging environments.

Extant Undergraduate Mentorship Research
Universities have provided faculty–student mentorship to foster student success in the STEM environment. Multiyear undergraduate research experiences have been shown to increase the number of PEER students who graduate from college and subsequently earn a PhD (Maton *et al.*, 2012, 2016; Estrada *et al.*, 2018). These studies have been essential to constructing the mentorship theoretical model and provide evidence that these experiences impact their science efficacy, identity, and values (Crisp and Cruz, 2009; Dolan and Johnson, 2009, 2010; Estrada *et al.*, 2018; Robnett *et al.*, 2018) However, independent research and course-based undergraduate research experiences as a form of mentorship are difficult to scale. At large-enrollment postsecondary institutions, the student:faculty ratios may be greater than 80:1 in certain majors (data from University of California, Santa Barbara [UCSB] Institutional Research, Planning, and Assessment), making it difficult, if not impossible, to provide all students in these programs with meaningful faculty mentorship. Therefore, near-peer mentorship programs that can be implemented at scale are an alternative to create these support structures. This approach generally involves pairing first-year undergraduates with third- or fourth-year students who have similar goals and backgrounds, with the mentor aiming to provide personalized guidance and insight into navigating the complex university system.

Despite implementation of near-peer mentorship programs across university campuses, rigorous empirical evidence regarding their effectiveness and mechanisms of action remains thin (as outlined in reviews by Jacobi, 1991; Crisp and Cruz, 2009; Gershenfeld 2014; Lane, 2020). In fact, the analysis of near-peer mentorship literature reviewed by Gershenfeld (2014) and Lane (2020) noted that the majority of mentorship studies provided no conclusive evidence that their programs had an impact on the desired outcomes; further, most lacked methodological rigor. This current standing of the field is likely due in part to the diversity of mentorship scholarship; in their review, Crisp and Cruz identified 50 distinct definitions of mentorship in published research studies. Further, these studies also varied in the potential outcome measures and methodology employed.

Of the studies available, some near-peer mentoring programs have demonstrated the potential to improve the transition to the university, promote positive student development, and improve academic performance (Johnson *et al.*, 1998; Hansford *et al.*, 2003; Rodger and Tremblay, 2003; Salinitri, 2005; Wilcox *et al.*, 2005; Sorrentino, 2006; Dearlove *et al.*, 2007; Nora and Crisp, 2007; Green, 2008; O'Brien *et al.*, 2012; Zaniewski and Reinholz, 2016). O'Brien *et al.* (2012) evaluated a 6-week near-peer mentoring program with first-year education students aimed at facilitating the transition to university by setting realistic goals, making students feel they were valued by the university, and facilitating the development of friendships. Based on results from pre- and postprogram questionnaires, mentees reported significantly less stress about coming to university and less worry about not belonging. These results suggest that the instrumental, psychosocial, and academic support provided by the mentors may influence mentee academic habits (e.g., academic support: providing insights into effective study habits) or perceptions of college (e.g., psychosocial support: emotional support in persisting through stressful setbacks), resulting in the observed improved academic performance and retention (Filippou *et al.*, 2015).

Although research into the effectiveness of near-peer mentorship provides encouraging results, these studies often have
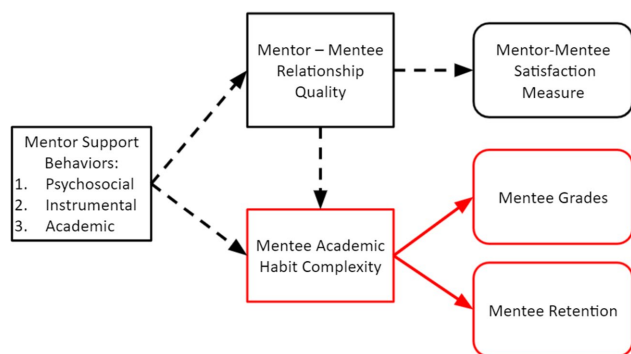
FIGURE 1. Conceptual model of BIOME peer mentorship for first-year biology majors. Squares represent theoretical elements of BIOME peer mentorship, while rounded boxes are measured outcomes. Red indicates framework outcomes measured in the present study. Dashed lines indicate hypothesized relationships based on extant literature.

TABLE 1. Descriptive data of the 2017–2019 biology student cohorts[a]

	BIOME	Total
Biology major students (n)	437	2920
Female	290 (66.36%)	1970 (67.47%)
PEERs (URM)	144 (32.95%)	1054 (36.1%)
EOP	157 (35.93%)	1011 (34.62%)
First generation	204 (46.7%)	1276 (43.7%)
Mean total SAT (total: 1600)	1473	1500 (1505 for non-BIOME cohort)

[a]Demographics of the Fall 2017–2019 course offerings. Description includes only declared biology majors with first-year standing. Percentages denote the composition of particular demographics of the declared biology majors present in the sections of the courses. Differences are not significant as assessed by multilevel logistic regression, in which cohort year is the random intercept variable.

small sample sizes, do not focus on STEM majors, or collect data from a single cohort. Further, very few peer mentorship publications link their results to the proposed mentorship conceptual model (Crisp and Cruz, 2009; Figure 1). To begin to answer these questions, we present our analysis of Biology Mentoring and Engagement (BIOME) near-peer mentorship on participating mentees. To accomplish this, we used a mixed-methods approach to identify and subsequently characterize a putative mechanism by which BIOME may promote mentee academic success and retention. By analyzing three cohort years, collectively comprising 437 first-year biology students, we are able to assess how the BIOME course, and its associated mentorship, influences a large number of diverse biology undergraduate students.

The Study Context

In terms of experimental design, selection of the biology major to assess the impacts of this approach has two strengths. First, biology is the largest undergraduate major on the university's campus, where annual matriculation is ~1100 students (Table 1). These cohorts include large populations of PEERs (>30%), educational opportunity program–eligible individuals (40%; EOP eligibility is conferred by parental socioeconomic status), and first-generation students (40%). This large, diverse student body enables us to assess the effectiveness of a scaled mentorship approach. Second, all biology majors must complete a yearlong, three-course series of General Chemistry (hereafter CHEM 1A in Fall, CHEM 1B in Winter, and CHEM 1C in Spring) before enrollment in the second year–level Introductory Biology courses. To remain on track in the major, students must earn at least a "C–" letter grade (1.7 on the 4.0 scale) in each course of the series. Approximately 25–30% of declared biology majors do not earn the minimum grade; in fact, students who earn less than a "C–" grade in the Fall offering of CHEM 1A are less likely to be retained to the Fall of their second year (Supplemental Table 1). Thus, the General Chemistry series presents an opportunity for us to assess the impact of course-based near-peer mentorship on a shared major academic challenge to student performance and retention.

BIOME Course and Mentorship Structures

The BIOME seminar course is co-taught (M.W. and E.G.-N.) in the Fall quarter as a 50-minute-long, one-unit weekly seminar course restricted to first-year biology majors. Each class meeting involves a short (estimated to take 20–45 minutes) reading and reflection assignment that forms the basis for the discussion in that week's 50-minute class seminar. Themes and topics were designed to foster the development of academic "soft skills," including growth mindset (Claro *et al.*, 2016), time management (George *et al.*, 2008; Krumrei-Mancuso *et al.*, 2013), and reflective and adaptive study skills (Dunlosky *et al.*, 2013; Supplemental Table 2). Before commencement of the Fall quarter, incoming first-year biology students were invited to enroll in the one-unit, pass/no pass seminar by email and at in-person orientation. Enrollment in BIOME was optional. Total BIOME enrollment was ~110–150 per cohort year, and students selected one of four sections that were offered at various times throughout the week. First-year students (hereafter "mentees") were randomly assigned to academically successful, upper-division biology students (hereafter "mentors") who mentored their groups for the entire 10-week quarter; by using weekly assignments to guide conversation, mentors were to provide psychosocial, academic, and instrumental support to their mentees (Figure 1; Supplemental Table 2). The mentors, seated at small discussion tables with four to six of their mentees, directed discussion at each table during class time while the instructors guided larger, class-wide discussions to launch or summarize seminar themes of each of the weekly assignments. The mentors received one unit of course credit for participating in a 10-week-long mentorship training seminar that met weekly for 50 minutes (led by M.W. and E.G.-N.). The mentor training course centers on weekly readings and discussions of techniques of effective mentorship; subsequently, participating mentors role-play situations to practice these approaches. Further, mentors were trained to facilitate conversation and provide feedback to mentees, rather than to act as content tutors in the BIOME class settings.

Summary and Research Questions

Here we describe the structure of the BIOME course and analyze the academic impact on mentored students in three academic cohorts (2017–2019; Table 1). Given the study context, wherein a considerable number of biology majors struggle

academically in General Chemistry, we sought to analyze whether BIOME promoted student academic success in these required courses while decreasing the short-term attrition of biology majors in their first year of study. Thus, to characterize the potential impacts of this peer mentorship course, we sought to address the following four questions:

- Does BIOME influence student adoption of more diverse academic habits?
- Does increased academic habit complexity result in improved grade performance in the major's required CHEM 1A course?
- Does BIOME improve student retention into the subsequent CHEM 1B course as a short-term indicator of retention?
- Does BIOME differ in impact on PEERs?

METHODS

To evaluate BIOME, we employed a mixed-methods study design that enabled us to describe, connect, and explain the outcomes of our study (Creswell and Plano Clark, 2011). We initiated this approach by using qualitative research methods to identify putative mechanisms by which the BIOME could be promoting student success. This approach identified the importance of implementing multiple academic habits for student success in CHEM 1A—the construct of student academic habit complexity. To characterize this construct, we developed a 22-item academic habit complexity survey instrument. Results from the qualitative approaches and instrument design are explained in depth in Clairmont (2020) and summarized in the Supplemental Material. We present here the quantitative approaches that enabled us to study the effect of BIOME on student academic habit complexity, student grades, and short-term retention. In what follows, we elaborate on each phase of the quantitative methodological approaches, the specific samples or data involved, and the results of our analyses (Table 2).

Analysis Plan

To analyze whether the BIOME program influenced mentee academic habit complexity, CHEM 1A grade, or short-term student retention, we built regression models to compare academic outcomes of BIOME mentees and their non-enrolled peers. Thus, BIOME enrollment, our key predictor of interest, was coded 1 if the student was enrolled in BIOME, and 0 otherwise. UCSB Institutional Research, Planning, and Assessment provided anonymized grade, retention, BIOME enrollment, and demographic data for all 2920 first-year biology students included in the three cohorts that make up this study. In what

follows, we describe the specific regression analyses performed to estimate the impact of BIOME on CHEM 1A grades, retention of students progressing into CHEM 1B, and whether various student subgroups benefit disproportionately from the program. Student ethnicity, parental income/education level, Scholastic Aptitude Test (SAT) scores, and high school grade point average (GPA), as well as UCSB course grade and enrollment, were provided in anonymous aggregate by UCSB Institutional Research, Planning, and Assessment. This study was conducted under the guidelines of the UCSB Office of Research Human Subjects approved Institutional Review Board (IRB) protocol no. 2-17-0610. Under this IRB, demographic data, course retention information, and final grade data were available for all first-year biology students.

Latent Regression Analysis of BIOME on Academic Habit Complexity

To study the effect of BIOME on student academic habits, we used a latent regression modeling strategy. A latent regression institutes the use of covariates in the Rasch model, in particular person-level covariates like belonging to a certain program. The use of the model is contingent on the initial data from the instrument of use fitting the Rasch model. In our latent regression model to predict academic habit complexity, the student-person covariate was whether a student belonged to BIOME or not. In this way, the effect of BIOME, with the indicator variable taking on 1 or 0, can be interpreted with the same unit and scale as the items of the academic habit complexity instrument, while accounting for measurement error—effectively carrying uncertainty forward and providing more accurate estimates of the effect of BIOME. For more on latent regression models or so-called explanatory item response theory (IRT) models, see Briggs (2008). The results from this analysis provide an estimated difference in academic habit complexity between students enrolled in BIOME or not.

Linear Regression Analysis of Student Grade

To estimate the treatment effect of BIOME on the key outcome of student grades for the first-quarter CHEM 1A, we used a linear regression model. Final student letter grades were transformed into a 4.0 GPA scale. The model used here involved all pretreatment/pre-BIOME covariates that are thought to account for incoming academic ability differences or potential to select into the program. Our first baseline regression model included variables collected by UCSB Institutional Research, Planning, and Assessment.

TABLE 2. Descriptions of data analyses

Outcome variable analyzed	Method	Years analyzed
Academic habit complexity	Novel items authored and fit to a Rasch model Measure validation using evidence from ethnographic observation, cognitive interviews, relationships to other variables, and other data	Two years: 2018–2019
Grade performance in General Chemistry 1A	linear regression. Checks of robustness -propensity score matching and subsequent multilevel linear regression Intraclass correlation for influence of clusters	Three years: 2017–2019 combined with cohort year as random intercept variable
Student retention to General Chemistry 1B	Multilevel logistic regression analysis, checks of robustness, propensity score, and subgroup (moderations) analysis	Three years: 2017–2019 combined with cohort year as random intercept variable

$$CHEM\widehat{GRADE}_s = \alpha + \beta_1 BIOME_s + \beta_2 HSGPA$$

$$+ \sum\nolimits_3^5 B_i SATSCORE_i + B_6 LogParentIncome$$

$$+ \beta_7 Ethnicity + \beta_8 ParentEd$$

where CHEMGRADE represents student grades in the Fall CHEM 1A course; α is the intercept, which represents the mean grade of first-year biology students not enrolled in BIOME; β_1 represents the coefficient for the effect of BIOME; and $BIOME_s$ is an indicator variable for BIOME membership. The additional variables represent potentially confounding variables that may mask the effect of BIOME if not adjusted for.

These variables included student high school GPAs, student SAT scores, the natural logarithm of parental income (because it was extremely skewed) and education status, and student ethnicity. Additionally, because a student's cohort may have an effect on student achievement, the year a student was admitted was considered. This model was run with three cohort years of data (2017, 2018, 2019). These years were included as both random effects in a mixed effects model and as dummy variables in a least-squares regression (single level), and neither yielded different estimates for BIOME (Gelman and Hill [2007] note that mixed-effects models are not likely to add much information to a study when there are fewer than five groups).

Additionally, to ease interpretation, math and verbal SAT scores were divided by 100. Further, because the writing portion of the SAT changed in format and scoring between admissions years used in the analyses, SAT writing scores were standardized within the admission cohort years. In the model above, SAT scores represent the three different SAT scores described (hence three different coefficients). Covariates and their estimates are listed in the Supplemental Material. A check for collinearity among predictors in all regressions was performed using a variance inflation factor (VIF) and generalized VIF (given multiple categorical predictors, we used GVIF; Fox, 2016). No VIF or GVIF values were greater than 4.5 in any model, which is less than the common rule of thumb, whereby a VIF of 10 should be considered possibly problematic (O'Brien, 2010). The covariates with the largest VIF values were the SAT writing and SAT verbal scores; however, removal of these covariates did not alter estimated effects of BIOME.

Logistic Regression Analysis of Short-Term Retention

We were interested in assessing whether BIOME influenced short-term student retention in the biology major. However, students generally do not declare themselves as "dropped out" (Rumberger, 2011; Uretsky and Henneberger, 2020). At our institution, it is not required that students officially change their majors to begin taking courses for other majors that are housed in the College of Letters and Sciences (or to stop taking courses required of the biology major, e.g., General Chemistry). Therefore, we use the approach of quantifying student progression through required pre-major courses, like CHEM 1A and 1B, as a proxy for student retention in the biology major.

To conduct this analysis, we studied the rate at which students took the second required General Chemistry course, CHEM 1B, on time as recommended by the institution. That is, if they took the second chemistry course in the quarter recom-

mended by the major, it is likely that the students' intent is to remain in the biology major. Therefore, we conducted a logistic regression in which students who took CHEM 1B in the Winter quarter of their first year were considered on time and retained over the short term. Hence, the logistic regression took the form, adjusting for demographic and background variables,

$$Logit\left(ON\ TI\widehat{ME}\ COURSE = 1\right) = \alpha + \beta_1 BIOME_s + \beta_2 HSGPA$$

$$+ \sum\nolimits_3^5 B_i SATSCORE_i + B_6 LogParenIncome + \beta_7 Ethnicity$$

$$+ \beta_8 ParentEd$$

where ON TIME COURSE is a dummy variable taking on the value of 1 if the student takes the CHEM 1B course in the Winter of the first year as recommended by the major, and 0 otherwise (CHEM 1B: 1 = on time; 0 = not on time, hereafter referred to as hindered). BIOME is a dummy variable representing whether a student took BIOME in the first year, and β_1 represents the treatment effect in BIOME in logits and can be converted to odds ratios via exponentiation. For the propensity score matching, due to the use of weighting, a quasibinomial link was used.

Student Subgroup Analysis

To understand whether BIOME had a differential impact depending on student subgroups, the analyses described above were modified to include an interaction term between student ethnicity and the treatment, BIOME. In each analysis described above, the models were modified to appear:

$$CHEM\widehat{GRADE}_s = \alpha + \beta_1 BIOME_S + \beta_2 HSGPA$$

$$+ \sum\nolimits_3^5 B_i SATSCORE_i + B_6 LogParenIncome$$

$$+ \beta_7 Ethnicity + \beta_8 ParentEd + \beta_9 BIOME_s * Ethnicity$$

$$Logit\left(ONTI\widehat{ME}\ COURSE = 1\right) = \alpha + \beta_1 BIOME_s + \beta_2 HSGPA$$

$$+ \sum\nolimits_3^5 B_i SATSCORE_i + B_6 LogParenIncome + \beta_7 Ethnicity$$

$$+ \beta_8 ParentEd + BIOME * Ethnicity$$

In these models, both Ethnicity and BIOME are categorical variables, so statistical significance implies a potential difference between any category and the reference group. In the model, the reference group used was white.

Propensity Score Matching

To conduct an analysis of the effect of BIOME participation on student academic outcomes that may account for the self-enrollment selection bias, we had to identify an appropriate control group of nonparticipating students. Therefore, we used propensity score matching to match each treatment student (in BIOME) to a non-BIOME member (matched sample student), simulating a randomized experiment but shifting the interpretation of the analysis in this portion (see Supplemental Material) to the causal effect of those who would be treated (the

average treatment effect on the treated). A concern with propensity score matching is that there are yet unmodeled determinants of BIOME membership including psychosocial factors that we have not measured. Further, there are concerns that propensity score matching itself can bias results in other ways (for a larger discussion, see King and Nielsen, 2019). These concerns are the motivation for presenting the baseline regression models in case propensity score methods are biasing our results. Thus, we present both propensity score matching and regression analyses side by side (as in Ou and Reynolds, 2010; Xu and Jaggers, 2011).

We used BIOME students' propensity scores to identify comparable non-BIOME control students, generating what we call the "matched sample" (see Supplemental Material for details). Our matching procedures primarily used genetic matching via the MatchIt package in R (Ho *et al.*, 2011). Using this algorithm, we were able to find a match for each treated unit (BIOME student) from the control units (non-BIOME student) that represents "closeness" of these units based on the distribution of propensity scores for those units being matched. The propensity score–matching procedure resulted in two groups (BIOME group $n = 393$ and matched sample control group $n = 344$). Because genetic matching with replacement was used, the procedure finds the best matches for each treatment, sometimes re-using control group students. A potential problem of 1:1 matching is that it can lead to potential poor matches in cases where two (or more) treatment units would match well to the same control unit (Diamond and Sekhon, 2005; Stuart, 2010). All covariates had standardized mean differences of less than $|0.10|$ following matching. We provide additional information concerning our qualitative and propensity score matching approaches in the Supplemental Material. The propensity score–matched populations were then analyzed via regression analyses as described. For propensity score matching models, collinearity is less of a concern, as coefficient estimates are not the primary focus.

RESULTS

Prior research and the mentorship conceptual model suggest that effective, high-quality peer mentorship results in improved student academic performance and retention. Therefore, we sought to evaluate the effects of the BIOME near-peer mentorship program on mentee outcomes by 1) assessing whether enrollment in BIOME increases student academic habit complexity, 2) characterizing whether academic habit complexity is related to grade performance in CHEM 1A, 3) determining whether participation in BIOME improves student retention into the subsequent CHEM 1B course, and 4) whether BIOME has disproportionate impacts on PEERs. The student population presented are from three cohort years (2017–2019) and comprise 2920 first-year biology majors (Table 1).

BIOME Influences Academic Habit Complexity

To assess whether BIOME influenced mentee academic habits differently than those of their non-enrolled peers, all first-year biology students were invited by email to complete the online academic habit complexity survey scale. Although the response rate was ~35–40% of the first-year biology cohorts, there were no significant demographic differences between BIOME and non-BIOME respondents. A latent regression was performed to estimate the difference in academic habit complexity between those enrolled in BIOME or not at the beginning and end of the Fall quarters (Table 3). The latent regression revealed a significant positive difference for the BIOME group in academic habit complexity of 0.3 logits; that is, BIOME participants showed a level of academic habit complexity roughly 5% higher than that of the comparison group at the end of the Fall quarter. To put this in context, the 0.3 logit difference on the academic habits construct can be applied to specific items. Although the logit difference between those in BIOME versus those who are not remains constant, the estimated probability that a student endorses an academic behavior varies by item difficulty (Supplemental Figure 1A) and person ability (Supplemental Figure 1B). For example, an average first-year biology student (who is not in BIOME) is estimated to have a 51% chance of reworking CHEM 1A problems from tests and quizzes, while an equivalent student in BIOME would have a 59% chance of reporting having reworked these problems. However, with an easier item, such as attending a campus tutoring session, the probability of endorsement is 52% versus 64% for non-BIOME students compared with BIOME mentees, respectively. Additionally, a multivariate regression provided evidence that academic habit complexity and grades in CHEM 1A were positively related: students earned an average of half a letter grade higher in chemistry for every four academic habits they reported, even after controlling for SAT math scores, BIOME enrollment, and high school GPA (Table 4; Supplemental Table 3; Clairmont, 2020).

BIOME Correlates with Improved Student Academic Performance

To assess the impact that BIOME has on student academic performance, we carried out a regression analysis in which student CHEM 1A final grade was regressed on the treatment variable, BIOME, as well as all demographic variables presented in the *Methods* section. Results from the unmatched sample are presented in Supplemental Table 4. After adjusting for confounders, including student demographics, we observed that CHEM 1A final grades for students in BIOME, holding values on all other pretreatment variables constant, were 0.19 GPA points higher than students in the matched sample ($p < 0.0001$; Supplemental Table 4, column 2). To assess whether this analysis was biased due to differences between the student populations enrolled in BIOME or not, we performed propensity score matching (Table 5). In the matched sample, the results should be interpreted as the effect of the BIOME on those who are enrolled in BIOME (the average treatment effect on the treated [ATT]). In this case, the increase in CHEM 1A final grade was 0.19 GPA ($p < 0.001$) points for a student in BIOME compared with students like them (matched on demographics and other confounding variables) in the matched sample (Table 5).

TABLE 3. Latent regression results with item responses to the academic habit complexity instrument as outcomes

	Logit units (SE)	*p* value
BIOME	0.29 (.14)	0.019
Item threshold parameters	22	
Number of parameters estimated	25	
Deviance	7639.06	

TABLE 4. Results of regressing final grades in CHEM 1A (in GPA points) on ability estimates from the academic habit complexity instrument

	Beta	SE	t	p
(Intercept)	−6.984	1.05	−6.654	<0.001
Academic habit complexity	0.176	0.057	3.107	0.002
Scientific notation format	−8.529e−7	4.885e−7	−1.746	0.082
SAT Math score	0.007	8.974e−4	7.858	<0.001
High school GPA	1.163	0.241	4.826	<0.001
Ethnicity				
Reference group: white				
Asian	0.018	0.132	0.136	0.892
PEER	−0.163	0.158	−1.031	0.304
Gender				
Reference group: female				
Male	0.114	0.126	0.91	0.364
First generation	−0.094	0.133	−0.704	0.482

BIOME Improves Student Retention

To assess whether BIOME promotes student progress through the prerequisite courses for the biology major, we analyzed student enrollment in CHEM 1B in the subsequent Winter quarter via logistic regression. Enrollment in CHEM 1B in the Winter quarter is considered "on time" and is required for a biology student to remain on track for the biology major. Students who do not take CHEM 1B in Winter quarter, due to repeating CHEM 1A or not enrolling in CHEM 1B, are termed "hindered." Without adjusting for covariates, students enrolled in BIOME were 75% more likely to take CHEM 1B on time than those not in BIOME (on-time BIOME = 331 students; hindered BIOME = 83; on-time not BIOME = 1670; hindered not BIOME = 733)

We regressed on-time course taking of CHEM 1B on various background demographics and the treatment of interest, BIOME membership, as described in the *Methods* section. In the unmatched sample, students in BIOME were much more likely to take CHEM 1B on time (Supplemental Table 5). In fact, holding all covariates constant, the odds of students enrolled in BIOME taking CHEM 1B on time are double those of students who are not enrolled in BIOME (odds ratio = 1.92, $p < 0.001$). To check these models, we used the matched sample described earlier to analyze on-time course taking. Here, holding all covariates constant, the odds of a student in BIOME taking CHEM 1B on time are 2.07 times those of students in the matched sample ($p < 0.001$; Table 6).

BIOME Impact on Student Subgroups

We were interested in assessing whether BIOME had a disproportionate benefit for PEER individuals, who leave the biology major at a higher rate than their non-minoritized peers. To analyze the potential impact of BIOME on students from different subgroups, we constructed regressions in which student chemistry grades in CHEM 1A and student on-time course taking of CHEM 1B were each regressed on the interaction between whether a student was in BIOME or not and the student's ethnicity. Using the unmatched sample, Supplemental Table 4 (column 2) shows that the coefficients representing the interactions between BIOME enrollment and student ethnicity are statistically insignificant in predicting student CHEM 1A grades. Results are similar in the matched sample (Table 5). The matched sample for six students of unknown ethnicity is quite

small, meaning there is too much uncertainty regarding the interaction terms for those individuals.

Correspondingly, Table 6 shows that the odds of taking CHEM 1B on time do not change as a function of subgroup and enrollment in BIOME. In other words, as for CHEM 1A grades, it does not appear that there is a difference in outcomes of on-time course taking for students in BIOME of one ethnicity over another. In the case of the matched sample, student subgroup estimates also have a lot of uncertainty about the estimates, meaning the interaction terms between subgroup and BIOME treatment are not significant ($p > 0.05$) and all parameter estimates are close to odds ratios of 1.

DISCUSSION

As part of a larger, mixed-methods study, we uncovered evidence that the BIOME course influences student academic behaviors (for qualitative analyses, see Clairmont, 2020; Supplemental Materials). Parallel to our qualitative evidence, the quantitative evidence is strongly suggestive of a positive association between participation in BIOME and increases in academic habit complexity. Taking a latent variable approach, we found a statistically significant difference between participants in the BIOME and non-participating first-year undergraduates (Table 4). Cross-sectional designs are limited in the extent to which they can demonstrate causation without further design features. However, as Onwuegbuzie and Leech (2004) have argued, mixed-methods designs are perhaps a more appropriate "gold standard" for research involving human subjects interacting with complex systems than more traditional randomized controlled trials that involve randomization only for the purpose of statistical inference. Given the potential benefits of participation in a mentoring program and lacking institutional support to conduct a randomized controlled trial, a detailed mixed-methods study is the most realistic method for exploring the potential causal mechanisms at work in this program (Creswell and Plano Clark, 2011).

Prior research characterizing near-peer mentorship programs have included postintervention outcome measures that center on academic performance (grades) or affective outcome measures (Johnson *et al.*, 1998; Hansford *et al.*, 2003; Sorrentino, 2006; Wilcox *et al.*, 2005; Dearlove *et al.*, 2007; Nora and Crisp, 2007; Green, 2008; O'Brien *et al.*, 2012; Leidenfrost

TABLE 5. Coefficient estimates using the propensity score–matched sample with CHEM1A GPA as the outcome variable of interest

	Beta	SE	*p* value
BIOME			
Reference group: not BIOME	—	—	—
BIOME	0.19	–0.06	0.002***
Admit quarter			
Reference group: 2017 cohort	—	—	—
2018 Cohort	–0.18	–0.09	0.06*
2019 Cohort	–0.26	–0.09	0.01***
Gender			
Reference group: female	—	—	—
Male	0.18	–0.07	0.005***
SAT Math score (divided by 100)	0.72	–0.06	0.00***
SAT Verbal score (divided by 100)	0.08	–0.1	0.47
Standardized SAT Writing score	0.09	–0.07	0.25
Ethnicit			
Reference group: white	—	—	—
Asian	–0.19	–0.08	0.02**
International	0.22	–0.13	0.09*
Unknown ethnicity	–0.24	–0.25	0.34
PEER	–0.16	–0.09	0.09*
High school GPA	0.72	–0.13	<0.0001***
Parent education			
Reference group: 2-year college graduate	—	—	—
4-year college graduate	0.06	–0.14	0.69
High school graduate	–0.17	–0.16	0.3
Missing information	0.15	–0.46	0.74
No high school	–0.23	–0.19	0.23
Postgraduate degree	–0.08	–0.14	0.59
Some college	–0.16	–0.16	0.33
Some high school	0.06	–0.18	0.74
Parent income (log scale)	0.005	–0.03	0.88
Constant	–5.7	–0.92	*p* < 0.0001***
Observations	730		
Log likelihood	–848.73		
Akaike information criterion		1739.47	

et al., 2014; Zaniewski and Reinholz, 2016). To our knowledge, the literature has yet to provide a rigorous characterization of the mechanisms that underpin how near-peer mentorship structures embedded in a structured course promote the observed positive outcomes. Therefore, our development and implementation of a tailored, novel academic habit complexity measure sheds light on a possible mechanism of course-based near-peer mentorship: mentors in a course context are able to meaningfully influence mentee adoption of effective approaches to challenging academic environments (Table 4). These results reinforce the mentorship conceptual model wherein the BIOME course promotes increased academic habit complexity of enrolled mentees (Figure 1).

In the BIOME mentorship framework, it is posited that, by providing students with effective mentorship supporting useful study techniques and academic soft skills, mentee use of certain academically beneficial behaviors will increase. In turn, employing a diverse suite of academic behaviors—increased academic habit complexity—will result in better grades. While academic performance may include student learning, subject matter retention (long or short term), and the ability to combine and critically assess ideas, we took a narrow view of academic performance as student performance in the course, operationalized via final grades. This decision represented, at least in the short term, the goals and values of students and faculty within the biology major. With this approach, we found evidence of a relationship between BIOME participation and student success in the concurrently offered, biology prerequisite course of CHEM 1A (Table 5; Supplemental Table 4). Using student grades in CHEM 1A, we analyzed the relationship between BIOME and student grades via linear regression. After adjusting for covariates, we saw that BIOME members' grades were 0.19 GPA points above those of students not in BIOME. In matched populations, our observed increase in CHEM 1A final grade was approximately the same 0.19 GPA points for students in BIOME compared with students like them. However, given that this is observational data, we cannot claim a causal relationship, because student participation in BIOME was optional. Further,

TABLE 6. Propensity score–matched sample regression results of CHEM 1B on-time course taking on BIOME

Variable	Outcome: on-time CHEM 1B Course Taking		
	Odds ratio	**SE**	**_p_ value**
BIOME			
Reference Group: not BIOME	—	—	—
BIOME	1.72	0.206	0.008
Admit quarter			
Reference group: 2017 cohort	—	—	—
2018 Cohort	0.65	0.315	0.2
2019 Cohort	0.61	0.304	0.11
Gender			
Reference group: female	—	—	—
Male	1.93	0.244	0.007
Ethnicity			
Reference group: white	—	—	—
Asian	0.96	0.298	>0.9
International	0.81	0.945	0.8
Unknown	0.39	1.25	0.4
PEER	0.65	0.299	0.14
SAT Math score (divided by 100)	3.29	0.231	<0.001
SAT Verbal score (divided by 100)	0.88	0.329	0.7
Standardized SAT Writing score	1.3	0.231	0.3
Parent income (on log scale)	1.07	0.065	0.3
High school GPA	3.34	0.416	0.004
Parent education			
2-year college graduate	—	—	—
4-year college graduate	1.15	0.484	0.8
HIGH school graduate	0.87	0.46	0.8
Missing	0.43	1.06	0.4
No high school	0.67	0.526	0.4
Postgraduate study	0.83	0.456	0.7
Some college	0.79	0.469	0.6
Some high school	1.24	0.521	0.7

the practical significance of this relationship is also difficult to glean. In the realm of letter grades at our institution, the difference between "C–" and a "C", a key cutoff for the biology majors (Supplemental Table 1), is ~0.3 GPA points. The observed grade increases, based on our analyses, are somewhere between two-thirds to four-fifths of this difference. The question is for whom and how this difference changes students' retention in the major or their perceived belongingness in the major, as this grade difference will not necessarily promote students into a new letter grade unless they are close to the grade boundaries.

Student integration into the university community is a key determinant of student retention (Tinto, 1993). Although near-peer mentorship programs have demonstrated promising qualitative results in influencing student perceptions of belonging (Yomtov *et al.*, 2017; Zaniewski and Reinholz, 2016; Lim *et al.*, 2017; Moschetti *et al.*, 2018) and intentions to persist (Solanki *et al.*, 2019), evidence of peer mentorship promoting STEM student retention beyond the program remains preliminary (Zaniewski and Reinholz, 2016). Given the results that BIOME mentees are approximately two times more likely to enroll in CHEM 1B, this result suggests that our peer mentorship course may promote student retention in the biology prerequisite course series.

Limitations

Assessing whether BIOME influenced student academic behaviors relied on the academic complexity measure (Clairmont, 2020; Supplemental Figure 1A). The constructs that make up the 22-item instrument arose from BIOME classroom ethnography combined with subsequent focus group discussions. Focus-group participants were first-year biology students who were enrolled in BIOME and CHEM 1A or CHEM 1A alone; importantly, students who were not enrolled in BIOME were able to identify the importance of the individual academic behaviors included in the instrument and academic success in CHEM 1A. Further, validity evidence was collected from students who were not included in study regression results presented and comprising individuals enrolled in BIOME and CHEM 1A or CHEM 1A alone. However, this approach is limiting, given that initial identification of constructs was heavily influenced by BIOME classroom characterization and analysis of student conversation within the course (see Supplemental Material); therefore, the instrument may be uniquely tailored to the academic context of students in our major and may not apply to other programs that center on delivering academic tutoring or other aspects of peer mentorship. This limitation, combined with the fact that the measure was only tested with the three cohorts of biology majors at our institution, may further limit the utility of

this instrument for measuring academic complexity with a broader population of students.

An additional challenge with our analyses and included regression models is that there may yet be unmeasured or unmodeled determinants of BIOME membership. In other words, while we would like to know whether BIOME causes certain improvements in student performance, there are various threats to the internal validity of this study (where internal validity is the extent to which causal claims can be based on the experimental or quasi-experimental design; Shadish *et al.*, 2002). Acknowledging these threats to internal validity means that we accept that there are possible alternative explanations or effects that cannot be eliminated by our quasi-experimental design. Because this study is based on observational data, we have identified various threats to internal validity. In the non–propensity score based models (Supplemental Tables 2 and 4), a threat arises in the form of bias being introduced by not adjusting for important covariates such as certain student propensities that may cause both membership in BIOME and student performance in the biology major. The only institutional data available for this study were collected upon matriculation of students (e.g., SAT score, high school GPA; all covariates included in Table 4 or Supplemental Table 4). This may mean that, for instance, that students in the program have knowledge, skills, or abilities that are different from those not in the program that have gone unmeasured. This also manifests itself in the propensity score matching models. Moving forward, increasing understanding of students who join or do not join the program may be helpful. However, our use of a mixed-methods approach of ethnographic methods, interviews, focus group discussions, and a specifically constructed instrument for academic habit complexity may help in understanding students' reasons for joining or not joining the program. Alternatively, if future peer mentorship programs are able to randomly assign students to participate as mentees, this would enable reduction of threats to internal validity while providing a route to inferring causal outcomes.

Although we were unable to randomly assign our participants, the two populations are demographically and academically comparable (Table 1); further, we performed propensity score matching in an effort to reduce any academic and demographic biases. Regardless of approach, we consistently observed that BIOME participants enrolled in CHEM 1B on time at greater rates than their non-BIOME peers (Table 6; Supplemental Table 5). Because part of our program design is to provide students with diversified and effective academic habits while informing student perceptions of the biology major (Figure 1), it will be important to analyze the longer-term outcomes of the cohorts of BIOME students to assess whether there are lasting impacts of this mentorship experience.

There is research showing that propensity score matching can actually increase bias in the estimation of causal effects in certain scenarios. For instance, King and Nielsen (2019) advises against using propensity scores to adjust for unobserved covariates or confounding. More specifically, a problem might arise under propensity score methods in which covariate imbalance is increased via propensity score matching. However, in the case of this study, while imbalance is not unimportant, the primary motivation for using propensity score matching was to achieve covariate overlap—that is, making sure that like students are compared (Gelman *et al.*, 2020).

In our propensity score model, we found a similar but slightly larger relationship between BIOME participation and grades. BIOME students in the matched sample scored approximately one-fifths of a GPA point higher in general chemistry than those not in BIOME (Table 5). As a reminder, this estimated effect is actually the average treatment effect on the treated, as some students who did not have matches were removed. It is of course possible that the propensity score model was mis-specified and either inflated or attenuated the estimated relationship between BIOME and grades. Regardless, this analysis provides evidence that BIOME may have an effect on student grades in CHEM 1A.

The structure of BIOME integrates metacognitive assignments over the 10-week quarter (Supplemental Table 2). These course materials are then combined with and delivered by near-peer mentors who are providing instrumental, academic, and psychosocial support (Figure 1; Supplemental Table 2, column 3). Therefore, disentangling how each of these elements of the program contributes to the observed mentee outcomes remains unknown. Thus, we cannot conclude that our observations are due to peer mentoring or course elements alone; rather, our outcomes are most likely due to the combination of mentorship within the structure of the BIOME course.

Future Research

In our context, BIOME promotes the diversification of mentee academic habits, improved academic performance in CHEM 1A (a required gateway STEM course), and mentee retention into the subsequent CHEM 1B course. Given the conceptual model presented in Figure 1, a critical next step will be to analyze activities or actions performed by the upper-division mentors to assess the combination of psychosocial, instrumental, and academic support behaviors that promote student academic behavior complexity. In a similar vein, assessing how these mentorship behaviors influence the fourth element of mentee–mentor relationship, quality (Rhodes *et al.*, 2017), will allow us to gain a greater insight into how this trait interacts with psychosocial, instrumental, and academic supports to influence mentee academic performance and retention. In addition to outlining what comprises effective peer mentorship, the conceptual model enables future research to use more sophisticated statistical approaches, including structural equation modeling or error-in-variable models, which can determine relationships between mentorship variables and academic outcomes while accounting for measurement uncertainty. From the perspective of the conceptual model, it would be beneficial to start explaining some of the mechanisms between inputs and outputs of the program with more depth—including more details about how specific mentor activities potentially effect mentees, which classroom activities are most useful in facilitating specific outcomes, and how students feel about these activities. Additionally, from the perspective of instrument construction, it will be important to follow up with students longitudinally to investigate whether students continue to diversify academic habits in future course work. Finally, it would be of methodological interest to consider model and item fit across students in BIOME/not in BIOME to further refine the academic habit complexity scale.

Extant literature on mentored research experiences has demonstrated evidence on the importance of aligning mentee–mentor demographics or values (Blake-Beard *et al.*, 2007, 2011; Terrion and Leonard, 2007; Hernandez *et al.*, 2016, 2017; Atkins *et al.*, 2020). Given the rapid proliferation of peer mentorship interventions across postsecondary educational settings, these programs provide a diversity of contexts that will enable dissection of how alignment of various mentor–mentee values and demographics influence the effectiveness of these approaches. Finally, characterizing whether peer mentorship influences mentee perceptions of their STEM majors, or their intent to persist, will enable greater understanding of whether these programs lead to changes in student affect that may result in longer-term impacts, including student retention.

ACKNOWLEDGMENTS

REFERENCES

Allen, T. D., & Eby, L. T. (2003). Relationship effectiveness for mentors: Factors associated with learning and quality. *Journal of Management*, *29*, 469–486. doi: 10.1016/S0149- 2063(03)00021-7

Allen, T. D., & Eby, L. T. (2007). Common bonds: An integrative perspective on mentoring. In Allen, T. D., & Eby, L. T. (Eds.), *The Blackwell handbook of mentoring: A multiple perspectives approach* (pp. 397–419). Oxford, UK: Blackwell. doi: 10.1111/b.9781405133739.2007.00024.x

Almatrafi, O., Johri, A., Rangwala, H., & Lester, J. (2017). *Retention and persistence among STEM students: A comparison of direct admit and transfer students across engineering and science Paper presented at 2017 ASEE Annual Conference & Exposition, Columbus, Ohio*.

Aryee, M. (2017). *College students' persistence and degree completion in science, technology, engineering, and mathematics (STEM): The role of non-cognitive attributes of self-efficacy, outcome expectations, and interest* (Doctoral dissertation). Seton Hall University dissertations and theses (ETDs), no. 2246. Retrieved May 16, 2021, from https://scholarship.shu.edu/dissertations/2246

Asai, D. J. (2020). Race matters. *Cell*, *181*, 754–757. https://doi.org/10.1016/j.cell.2020.03.044

Atkins, K., Dougan, B. M., Dromgold-Sermen, M. S., Potter, H., Sathy, V., & Panter, A. T. (2020). "Looking at myself in the future": How mentoring shapes scientific identity for STEM students from underrepresented groups. *International Journal of STEM Education*, *7*(1), 42. doi: 10.1186/s40594-020-00242-3 Epub 2020 Aug 18. PMID: 32850287; PMCID: PMC7437647.

Austin, A. E. (2002). Preparing the next generation of faculty. *Journal of Higher Education*, *73*, 94–122. doi: 10.1353/jhe.2002.0001

Barge, S., & Gehlbach, H. (2012). Using the theory of satisficing to evaluate the quality of survey data. *Research in Higher Education*, *53*(2), 182–200. https://doi.org/10.1007/s11162-011-9251-2

Blake-Beard, S., Bayne, M., Crosby, F., & Muller, C. (2011). Matching by race and gender in mentoring relationships: Keeping our eyes on the prize. *Journal of Social Issues*, *67*, 622–643.

Blake-Beard, S. D., Murrell, A., & Thomas, D. (2007). Unfinished business: The impact of race on understanding mentoring relationships. In Ragins, B. R., & Kram, K. E. (Eds.), *The handbook of mentoring at work: Theory, research, and practice* (pp. 223–247). Thousand Oaks, CA: Sage.

Blinn-Pike, L. (2007). The benefits associated with youth mentoring relationships. In Allen, T. D., & Eby, L. T. (Eds.), *The Blackwell handbook of mentoring: A multiple perspectives approach* (pp. 165–187). Oxford, UK: Blackwell.

Briggs, D. C. (2008). Using explanatory item response models to analyze group differences in science achievement. *Applied Measurement in Education*, *21*(2), 89–118, doi: 10.1080/08957340801926086

Chang, M., Sharkness, J., Hurtado, S., & Newman, C. (2014). What matters in college for retaining aspiring scientists and engineers from underrepresented racial groups. *Journal of Research in Science Teaching*, *51*, 555–580.

Chemers, M. M., Hu, L., & Garcia, B. F. (2001). Academic self-efficacy and first year college student performance and adjustment. *Journal of Educational Psychology*, *93*, 55–64. doi: 10.1037//0022-0663.93.1.55

Clairmont, A. (2020). *A discovery-based workflow for educational measurement* (Doctoral dissertation). University of California, Santa Barbara, CA: ProQuest. https://www.proquest.com/openview/c65715106aec663ae7deca18bb05230a/1?pq-origsite=gscholar&cbl=18750&diss=y

Claro, S., Paunesku, D., & Dweck, C. S. (2016). Mindset tempers effects of poverty on achievement. *Proceedings of the National Academy of Sciences USA*, *113*(31), 8664–8668; doi: 10.1073/pnas.1608207113

Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.). Thousand Oaks,CA: Sage.

Crisp, G., & Cruz, I. (2009). Mentoring college students: A critical review of the literature between 1990 and 2007. *Research in Higher Education*, *50*(6), 525–545.

Dearlove, J., Farrell, H., Handa, N., & Pastore, C. (2007). The evolution of peer mentoring at the University of Western Sydney. *Journal of the Australia and New Zealand Student Services Association*, *29*, 21–35.

Diamond, A., & Sekhon, J. (2005). Genetic matching for estimating causal effects: A new method of achieving balance in observational studies.

Dolan, E.& .L & Johnson, D. (2009). Toward a holistic view of undergraduate research experiences: An exploratory study of impact on graduate/postdoctoral mentors. *Journal of Science Education and Technology*, *18*, 487–500.

Dolan, E. L., & Johnson, D. (2010). The undergraduate-postgraduate-faculty triad: Unique functions and tensions associated with undergraduate research experiences at research universities. *CBE—Life Sciences Education*, *9*(4), 543–553. https://doi.org/10.1187/cbe.10-03-0052

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, *14*(1), 4–58. https://doi.org/10.1177/1529100612453266

Eagan, M. K., Stolzenberg, E. B., Zimmerman, H. B., Aragon, M. C., Whang Sayson, H., & Rios-Aguilar, C. (2017). *The American freshman: National norms Fall 2016*. Los Angeles: Higher Education Research Institute, UCLA.

Eby, L. T. D. T., Allen, T. D., Hoffman, B. J., Baranik, L. E., Sauer, J. B., Baldwin, S. … & Evans, S. C.(2013). An interdisciplinary meta-analysis of the potential antecedents, correlates, and consequences of protege perceptions of mentoring. *Psychological Bulletin*, *139*(2), 441–476.

Estrada, M., Hernandez, P., & Schultz, W. (2018). A longitudinal study of how quality mentorship and research experience integrate underrepresented minorities into STEM careers. *CBE—Life Sciences Education*, *17*(1), ar9.

Estrada, M., Woodcock, A., Hernandez, P. R., & Schultz, P. W. (2011). Toward a model of social influence that explains minority student integration into the scientific community. *Journal of Educational Psychology*, *103*(1), 206–222. https://doi.org/10.1037/a0020743

Fabert, N. (2014). *Growth mindset training to increase women's self-efficacy in science and engineering: A randomized-controlled trial* (Doctoral dissertation by Arizona State University). Retrieved May 16, 2021, from https://repository.asu.edu/attachments/137308/content/Fabert_asu_0010E_14151.pdf

Filippou, J., Cheong, C., & Cheong, F. (2015). Designing persuasive systems to influence learning: Modelling the impact of study habits on academic performance. In *PACIS 2015 Proceedings of Pacific Asia Conference on Information Systems (PACIS), Singapore held form July 6–9, 2015*.

Flaxman, E., Ascher, C., & Harrington, C. (1988). *Youth mentoring: Programs and practices.* New York: ERIC Clearinghouse on Urban Education, Institute for Minority Education, Teachers College, Columbia University.

Fox, J. (2016). *Applied Regression Analysis and Generalized Linear Models* (3rd ed.). Thousand Oaks, CA: Sage.

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge: Cambridge University Press.

Gelman, A., Hill, J., & Vehtari, A. (2020). *Regression and other stories*. Cambridge: Cambridge University Press.

George, D., Dixon, S., Stansal, E., Lund Gelb, S., & Pheri, T. (2008). Time diary and questionnaire assessment of factors associated with academic and personal success among university undergraduates, *Journal of American College Health*, *56*(6), 706–715, doi: 10.3200/JACH.56.6.706-715

Gershenfeld, S. (2014). A review of undergraduate mentoring programs. *Review of Educational Research*, *84*, 365–391.

Green, A. (2008). Straddling the gap: How second-year peers empower first-year students to participate in a community of independent learning by means of the educative. *Research in Post-compulsory Education*, *13*, 241–249.

Griffin, A. M., & Romm, J.(2008). *Exploring the evidence*, Vol. 4, *Reporting research on first-year seminars*. Columbia: National Resource Center for the First-Year Experience and Students in Transition, University of South Carolina.

Griffith, A. L. (2010). *Persistence of women and minorities in STEM field majors: Is it the school that matters?* (Cornell University, School of Industrial and Labor Relations working paper). Cornell Higher Education Research Institute, Cornell University. Retrieved May 16, 2021, from http://digitalcommons.ilr.cornell.edu/workingpapers/122

Hall, R., & Sandler, B. (1982). *The classroom climate: A chilly one for women?*. Washington, DC: Project on the Status of Education of Women, Association of American Colleges.

Hansford, B., Tennent, L., & Ehrich, L. (2003). Educational mentoring: Is it worth the effort? *Education Research and Perspectives*, *30*, 42–75.

Hernandez, P. R., Estrada, M., Woodcock, A., & Schultz, P. W. (2016). Protégé perceptions of high mentorship quality depend on shared values more than on demographic match. *Journal of Experimental Education*, *85*(3), 450–468.

Hernandez, P. R., Bloodhart, B., Barnes, R. T., Adams, A. S., Clinton, S. M., Pollack, I. … & Fischer, E. V. (2017). Promoting professional identity, motivation, and persistence: Benefits of an informal mentoring program for female undergraduate students. *PLoS ONE*, *12*(11), e0187531. https://doi.org/10.1371/journal.pone.0187531

Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2011). MatchIt: Nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software*, *42*(8), 1–28. Retrieved May 16, 2021, from www.jstatsoft.org/v42/i08

Hoffman, M., Richmond, J., Morrow, J., & Salomone, K. (2002). Investigating "sense of belonging" in first-year college students. *Journal of College Student Retention: Research, Theory & Practice*, *4*, 227–256. doi: 10.2190/dryc-cxq9-jq8v-ht4v

Jacobi, M. (1991). Mentoring and undergraduate academic success: A review of the literature. *Review of Educational Research*, *61*, 505–532. doi: 10.3102/00346543061004505

Johnson, D. W., Johnson, R. T., & Smith, K. A. (1998). Cooperative learning returns to college what evidence is there that it works. *The Magazine of Higher Learning*, *30*(4), 26–35.

Johnson, W. B., Rose, G., & Schlosser, L. Z. (2007). Student-faculty mentoring: Theoretical and methodological issues. In Allen, T. D., & Eby, L. T. (Eds.), *The Blackwell handbook of mentoring: A multiple perspectives approach* (pp. 49–69). Oxford, UK: Blackwell. doi: 10.1002/9780470691960.ch4

King, G., & Nielsen, R. (2019). Why propensity scores should not be used for matching. *Political Analysis*, *27*(4), 435–454

Kram, K. (1985). *Mentoring at work*. Glenview, IL: Scott Foresman.

Kram, K. (1988). *Mentoring at work: Developmental relationships in organizational life* (2nd ed.). Lanham, MD: University Press of America.

Krumrei-Mancuso, E. J., Newton, F. B., Kim, E., & Wilcox, D. (2013). Psychosocial factors predicting first-year college student success. *Journal of College Student Development*, *54*(3), 247–266. doi: 10.1353/csd.2013.0034

Lane, S. R. (2020). Addressing the stressful first year in college: Could peer mentoring be a critical strategy? *Journal of College Student Retention: Research, Theory & Practice*, *22*(3), 481–496. https://doi.org/10.1177/1521025118773319

Leidenfrost, B., Strassnig, B., Schütz, M., Carbon, C. C., & Schabmann, A. (2014). The impact of peer mentoring on mentee academic performance: Is any mentoring style better than no mentoring at all? *International Journal of Teaching and Learning in Higher Education*, *26*, 102–111. doi: 10.1080/13611267.2011.597122

Lim, J. H., MacLeod, B. P., Tkacik, P. T., & Dika, S. L. (2017). Peer mentoring in engineering: (Un)shared experience of undergraduate peer mentors and mentees. *Mentoring & Tutoring: Partnership in Learning*, *25*(4), 395–416, doi: 10.1080/13611267.2017.1403628

Maton, K. I., Beason, T. S., Godsay, S., Domingo, M., Bailey, T., Sun, S., & Hrabowski, F., III. (2016). Outcomes and processes in the Meyerhoff Scholars Program: STEM PhD completion, sense of community, perceived program benefit, science identity, and research self-efficacy. *Life Sciences Education*, *15*(3), ar48.

Maton, K. I., Pollard, S. A., McDougall Weise, T. V., & Hrabowski, F. A. (2012). Meyerhoff Scholars Program: A strengths-based, institution-wide approach to increasing diversity in science, technology, engineering, and mathematics. *Mount Sinai Journal of Medicine: A Journal of Translational and Personalized Medicine*, *79*, 610–623.

Miller, A. (2002). *Mentoring students & young people: A handbook of effective practice*. London, UK: Kogan Page Limited.

Morgan, S. L., & Winship, C. (2015). *Counterfactuals and causal inference*. New York, NY: Cambridge University Press.

Moschetti, R. V., Plunkett, S. W., Efrat, R., & Yomtov, D. (2018). Peer mentoring as social capital for Latina/o college students at a Hispanic-serving institution. *Journal of Hispanic Higher Education*, *17*(4), 375–392. https://doi.org/10.1177/1538192717702949

Nakkula, M. J., & Harris, J. T. (2005). Assessment of mentoring relation- ships. In . DuBois, D. L., & Karcher, M. J. (Eds.), *Handbook of youth mentoring* (pp. 100–117). Thousand Oaks, CA: Sage.

National Academies of Sciences, Engineering, and Medicine (NASEM). (2018). *Indicators for monitoring undergraduate STEM education*. Washington, DC: National Academies Press. https://doi.org/10.17226/24943

NASEM. (2019). *The science of effective mentorship in STEMM*. Washington, DC: National Academies Press. https://doi.org/10.17226/25568

National Research Council (NRC). (2009). *A new biology for the 21st century*. Washington, DC: National Academies Press.

NRC. (2015). *Reaching students: What research says about effective instruction in undergraduate science and engineering* (p. 256). Washington, DC: National Academies Press.

Nora, A., & Crisp, G. (2007). Mentoring students: Conceptualizing and validating the multi-dimensions of a support system. *Journal of College Student Retention: Research, Theory and Practice*, *9*(3), 337–356. ISSN 1521-0251, 1541-4167.

O'Brien, J. L. (2010). *Structural empowerment, psychological empowerment and burnout in registered staff nurses working in outpatient dialysis centers* (Doctoral dissertation, Rutgers University-Graduate School-Newark).

O'Brien, M., Llamas, M., & Stevens, E. (2012). Lessons learned from four years of peer mentoring in a tiered group program within education. *Journal of the Australian and New Zealand Student Services Association*, *40*(1), 7–15.

Onwuegbuzie, A. J., & Leech, N. L. (2004). Enhancing the interpretation of "significant" findings: The role of mixed methods research. *Qualitative Report*, *9*(4), 770–792.

Ostrove, J. M., & Long, S. M. (2007). Social class and belonging: Implications for college adjustment. *Review of Higher Education*, *30*, 363–389.

Ou, S.-R., & Reynolds, A. J. (2010). Grade retention, postsecondary education, and public aid receipt. *Educational Evaluation and Policy Analysis*, *32*(1), 118–139. https://doi.org/10.3102/0162373709354334

Patton, M. Q. (2008). *Utilization-focused evaluation*. Thousand Oaks, CA: Sage.

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.

Rasch, G. (1960). Probabilistic models for some intelligence and attainment tests. Danmarks Paedagogiske Institut.

Rhodes, J. E. (2002). *Stand by me: The risks and rewards of mentoring today's youth*. Cambridge, MA: Harvard University Press.

Rhodes, J. E. (2005). A model of youth mentoring. In DuBois, D. L., & Karcher, M. J. (Eds.), *Handbook of youth mentoring* (pp. 30–43). Thousand Oaks, CA: Sage.

Rhodes, J. E., Schwartz, S. E. O., Willis, M. M., & Wu, M. B. (2017). Validating a mentoring relationship quality scale: Does match strength predict match length? *Youth & Society*, *49*, 415–437. https://doi.org/10.1177/0044118X14531604

Riegle-Crumb, C., King, B., & Irizarry, Y. (2019). Does STEM stand out? Examining racial/ethnic gaps in persistence across postsecondary fields. *Educational Researcher*, *48*, 133–144. https://doi.org/10.3102/0013189X19849410

Robnett, R. D., Nelson, . A., Zurbriggen, E. L., Crosby, F. J., & Chemers, M. M. (2018). Research mentoring and scientist identity: Insights from undergraduates and their mentors. *International Journal of STEM Education*, *5*, ar41. https://doi.org/10.1186/s40594-018-0139-y

Rodger, S., & Tremblay, P. F. (2003). The effects of a peer mentoring program on academic success among first year university students. *Canadian Journal of Higher Education*, *33*(5), 1–18.

Rumberger, R. W. (2011). *Dropping out*. Cambridge, MA: Harvard University Press.

Salinitri, G. (2005). The effects of formal mentoring on the retention rates for first-year, low achieving students. *Canadian Journal of Education*, *28*, 853–873.

Scandura, T. A. (1992). Mentorship and career mobility: An empirical investigation. *Journal of Organizational Behavior*, *13*, 169–174. doi: 10.1002/job.4030130206

Schockett, M. R., & Haring-Hidore, M. (1985). Factor analytic support for psychosocial and vocational mentoring functions. *Psychological Reports*, *57*, 627–630.

Seymour, E., & Hewitt, N. M. (1997). *Taking about leaving: Why undergraduates leave the sciences*. Boulder, CO: Westview.

Shadish, R. W., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.

Solanki, S. M., McPartlan, P., Xu, D., & Sato, B. K. (2019). Success with EASE: Who benefits from a learning community? *PLoS ONE*, *14*(3), e0213827. doi: https://doi.org/10.1371/journal.pone.0213827

Sorrentino, D. M. (2006). The SEEK mentoring program: An application of the goal-setting theory. *Journal of College Student Retention Research Theory and Practice*, *8*(2), 241–250. doi: 10.2190/7D9T-D30Y-N9T0-8DWL

Spencer, R. (2007). Naturally occurring mentoring relationships involving youth. In Allen, T. D., & Eby, L. T. (Eds.), *The Blackwell handbook of mentoring: A multiple perspectives approach* (pp. 99–117). Oxford, UK: Blackwell. doi: 10.1002/9780470691960.ch7

Stains, M., Harshman, J., Barker, M. K., Chasteen, S. V., Cole, R., DeChenne-Peters, S. E., ... & Young, A. M. (2018). Anatomy of STEM teaching in North American universities. *Science*, *359*(6383), 1468–1470. https://doi.org/10.1126/science.aap8892 Medline, Google Scholar

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science: A Review Journal of the Institute of Mathematical Statistics*, *25*(1), 1.

Tenenbaum, H. R., Crosby, F. J., & Gliner, M. D. (2001). Mentoring relationships in graduate school. *Journal of Vocational Behavior*, *59*, 326–341. doi: 10.1006/jvbe.2001.1804

Terrion, J. L. (2012). Student peer mentors as a navigational resource in higher education. In Fletcher, S. J., & Mullen, C. A. (Eds.), *The Sage handbook of mentoring and coaching in education* (pp. 383–397). London, UK: Sage. doi: 10.4135/9781446247549.n26

Terrion, J. L., & Leonard, D. (2007). A taxonomy of the characteristics of student peer mentors in higher education: Findings from a literature review, *Mentoring & Tutoring: Partnership in Learning*, *15*(2), 149–164, doi: 10.1080/13611260601086311

Tinto, V. (1988). Stages of student departure: Reflections on the longitudinal character of student leaving. *Journal of Higher Education*, *59*(4), 438–455 https://doi.org/10.1080/00221546.1988.11780199

Tinto, V. (1993). *Leaving college: Rethinking the causes and cures of student attrition* (2nd ed.). Chicago: University of Chicago Press.

Tinto, V. (2006). Enhancing student persistence: Lessons learned in the United States. *Aná Psicológica*, *24*(1), 7–13.

Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*. New York, NY: Cambridge University Press.

Uretsky, M. C., & Henneberger, A. K. (2020). Fifth year persisters: High school, college, and early career outcomes for persisting non-graduates. *Children and Youth Services Review*, *115*, 105058.

Verplanken, B., & Aarts, H. (1999). Habit, attitude, and planned behaviour: Is habit an empty construct or an interesting case of goal-directed automaticity? *European Review of Social Psychology*, *10*(1), 101–134. https://doi.org/10.1080/14792779943000035

Wilcox, P., Winn, S., & Fyvie-Gauld, M. (2005). "It was nothing to do with the university, it was just the people": The role of social support in the first-year experience of higher education. *Studies in Higher Education*, *30*(6), 707–722.

Wilton, M., Gonzalez-Nino, E., McPartlan, P., Terner, Z., Christoffersen, R. E., & Rothman, J. H. (2019). Improving academic performance, belonging, and retention through increasing structure of an introductory biology course. *CBE—Life Sciences Education*, *18*(4), ar53.

Xu, D., & Jaggars, S. S. (2011). The effectiveness of distance education across Virginia's community colleges: Evidence from introductory college-level math and English courses. *Educational Evaluation and Policy Analysis*, *33*(3), 360–377.

Yomtov, D., Plunkett, S. W., Efrat, R., & Marin, A. G. (2017). Can peer mentors improve first-year experiences of university students. *Journal of College Student Retention: Research, Theory & Practice*, *19*(1), 25–44..

Zaniewski, A., & Reinholz, D. (2016). Increasing STEM success: A near-peer mentoring program in the physical sciences. *Journal of STEM Education*, *3*, ar14. doi 10.1186/s40594-016-0043-2