Published in final edited form as:
Educ Eval Policy Anal. 2011 March 1; 33(1): . doi:10.3102/0162373710392371.

Patching the Pipeline: Reducing Educational Disparities in the Sciences Through Minority Training Programs

P. Wesley Schultz,
California State University-San Marcos

Paul R. Hernandez,
University of Connecticut

Anna Woodcock,
Purdue University

Mica Estrada,
California State University-San Marcos

Randie C. Chance,
California State University-San Marcos

Maria Aguilar, and
California State University-San Marcos

Richard T. Serpe
Kent State University

Abstract

For more than 40 years, there has been a concerted national effort to promote diversity among the scientific research community. Yet given the persistent national-level disparity in educational achievements of students from various ethnic and racial groups, the efficacy of these programs has come into question. The current study reports results from a longitudinal study of students supported by a national National Institutes of Health–funded minority training program, and a propensity score matched control. Growth curve analyses using Hierarchical Linear Modeling show that students supported by Research Initiative for Science Excellence were more likely to persist in their intentions to pursue a scientific research career. In addition, growth curve analyses indicate that undergraduate research experience, but not having a mentor, predicted student persistence in science.

Keywords

science training programs; broadening participation; STEM; HLM

The scientific community benefits from diversity. Innovation, creativity, and novel discoveries are accelerated by a diversity of ideas and perspectives. While the scientific method provides a crucible for testing and validating these ideas, a diverse research

Address correspondences to P. Wesley Schultz, Department of Psychology, California State University, San Marcos, CA 92078, wschultz@csusm.edu.

Declaration of Conflicting Interests
The authors declared no potential conflicts of interests with respect to the authorship and/or publication of this article.

community with many perspectives affords a rich environment for new theories and hypotheses. In addition, a diverse research community can serve an important social role in society, with scientists serving as role models and encouraging underrepresented students to pursue academic interests across all disciplines. To this end, national-level institutions and policies have actively worked to promote a diverse research community. These efforts have been most notable in the STEM (Science, Technology, Engineering, and Mathematics) disciplines, with numerous programs across a range of educational levels and with support from a multitude of organizations and agencies (National Institutes of Health [NIH], 2005; National Science Foundation [NSF], 2000). Yet despite these efforts, the diversity of the scientific research community fails to reflect the broader U.S. population, and there are clear educational disparities across a number of demographic variables, including gender, race/ethnicity, and economic background. In this article, we provide new data examining the processes and outcomes of one of the largest and longest running minority science training programs in the country.

National data going back more than 40 years have documented sizable educational disparities across racial and ethnic groups in the United States (Cook & Córdova, 2006; Hanson, 2009; Johnson, 1997; National Center for Education Statistics [NCES], 2005). These data have consistently shown that students from Hispanic/Latino, African American, American Indian, and Pacific Islander ethnic/racial groups are underrepresented at all levels of higher education, and especially in science-related fields and careers. Despite making up over 20% of the U.S. adult population, individuals from these minority groups are underrepresented in the sciences at all levels of higher education, from undergraduate majors to graduate programs to university faculty positions (Orfield, Losen, Wald, & Swanson, 2004). In the areas of biology, psychology, and biochemistry (the scientific disciplines related to NIH-funded biomedical and behavioral research), the number of minority students who enter research careers is particularly small. In 2005, just 14% of bachelor's degrees, 10% of master's degrees, and 8% of doctoral degrees in the biological and life sciences were awarded to African Americans, Hispanics, and Native Americans combined (DePass & Chubin, 2008). Although these numbers reflect an increase over the prior decade, it is clear that minorities are still being lost at every step of the pathway toward a research career (NSF, 2000). There simply has not been a large enough increase in the retention of minority students in graduate and undergraduate degree programs to facilitate a more equitable representation in either academic or private-industry scientific research careers.

Prior studies of educational programs targeting underrepresented minority students (URMs) have focused on dichotomous outcomes—for example, graduation rates, applications to graduate programs, or choosing a science-related research career. While focusing on such outcomes is certainly warranted, we believe that it is equally important to understand the social and psychological processes that underlie these distal outcomes. Our focus in this article is on the persistence of students' intention to pursue a research career. Previous studies have shown the importance of intention in understanding longer term distal outcomes among university students (Cabrera, Nora, & Castaneda, 1993; Hausmann, Schofield, & Woods, 2007). And studies of intention have shown that upon entering college, a sizeable percentage of URM students have an interest in scientific careers and that many intend to pursue a research career (Hueftle, Rakow, & Welch, 1983; Hurtado et al., 2008). But over the course of their undergraduate studies, these intentions to persist in the sciences fluctuate. Focusing on student intentions to persist allows us to explore the psychological process that are linked with educational achievement, to statistically model changes over time in these intentions, and to connect program-level educational experiences with changes in intention and ultimately with academic achievement. In addition, focusing on intention allows for the multiple interests and career aspirations that are common among undergraduate students, and it allows for a more fluid perspective on student academic and career pathways during this time. Although sustaining interests in science is important, an equally strong case can be made for cultivating student interests, particularly in introductory or gateway courses to STEM disciplines.

Closing the Gap

For over 40 years, the implementation of educational intervention/training programs in schools and colleges has been one of the most common methods employed to help close this gap in educational achievement (Maton & Hrabowski, 2004). Thousands of these programs are being run on campuses throughout the United States each year. Funding for these programs comes from a range of sources, both public and private. Although there is no one source quantifying the number of programs implemented to address this gap, or the money spent on them, an indication of the scope of this expenditure is reflected in the billions of dollars spent on college readiness and science training programs. In 2004, the federal government spent $2.8 billion to increase the number of students in STEM fields. This money was spent on more than 200 different programs implemented in all 50 of the U.S. states (U.S. General Accounting Office, 2005). As specific examples, the NIH and NSF each funded a range of programs to support underrepresented students in the health sciences and STEM disciplines.

Minority educational intervention and training programs vary widely in their approach and the services they offer. Too numerous to list in full, the services offered by programs typically include one or more of the following: mentoring, stipends, test preparation, tutoring and specific skills training, college or graduate school preparation and exposure, research opportunities, enrichment programs and activities, supplemental instruction, and summer training programs. For example, the NSF-sponsored Louis Stokes Alliances for Minority Participation (LSAMP or AMP) initiative supports programs that provide students with a variety of enriching experiences, such as tutoring, summer research experiences, and direct financial support. Another example is the National Institutes of General Medical Sciences–sponsored Minority Access to Research Careers (MARC) program, which supports a variety of enriching experiences but requires participation in research activities at the students' home campus during the academic year and, furthermore, requires that the students participate in summer research experiences at research-intensive universities outside of their home institution. The proximal goals of both the AMP and MARC programs are to increase the quantity and quality of minority students completing their baccalaureate in a STEM discipline. Their intermediate goal is to increase the number of minority students pursuing graduate degrees in a STEM disciplines and ultimately to increase the number of minority STEM scientists. These goals imply persistence in the academic pipeline.

Although well intentioned, do these programs work? Are programs aimed at increasing the academic success of minority students meeting their goals? Despite the proliferation of minority educational intervention/training programs and the substantial funds spent on them, empirical evaluations are sparse and often methodologically unsound (Collea, 1990; Harrell & Forney, 2003; Lam, Mawasha, Doverspike, McClain, & Vesalo, 2000). A review of minority mentoring programs in higher education (Haring, 1999) concluded that they are notoriously difficult to evaluate, and in aggregate, they have failed to yield substantial growth in the numbers of minorities obtaining college degrees (Haring, 1999).

Although the tide is gradually turning, educational training programs, including large federally funded programs, have required very little in terms of evaluation. Evaluation plans and funding are written into each proposal, but more often than not, evaluation entails simply reporting the numbers of students served and the activities or programs carried out (Mervis, 2003). This is often supplemented by an anecdotal narrative about the program, or its elements, and their assumed link with outcomes. No systematic evaluation of these programs using established social scientific research methods is required, and as a result, there is little evidence about the efficacy of these programs (Haring, 1999; Jun & Tierney, 1999). In addition to a lack of summative evaluations, there is also no clear understanding of *why* a program succeeds or the essential elements of an effective educational intervention. There are also little data examining the types of students that are most likely to benefit from educational interventions. Indeed, given the stringent selection criteria used to admit students into many of these programs, it seems likely that they were already well on their way to succeeding prior to participation.

There is a lack of data demonstrating the efficacy of the thousands of intervention programs intended to help minority students succeed in higher education. Many intervention programs operate under the *assumption* that they are effective, although there is little data to support such a conclusion (see also NIH, 2005). This point was made emphatically in a 2006 *Science* publication based on a National Academy of Science report focused on science training programs for URM students, lamenting the lack of quality evaluation data (Mervis, 2006).

The missing piece of this equation is an empirical body of data, built upon rigorous evaluation of existing programs, that can serve as reliable counsel to program directors and funding agencies (cf., Bickman, 2000; Campbell, 1988). Although clusters of empirical data suggest the efficacy of a number of the key components of these programs (mentoring, research experience, and the provision of financial support, for example) or evaluations of individual campus programs (Maton, Hrabowski, Ozlemir, & Wimms, 2008), there is no multisite study available in the public domain that explores the impact of minority training science programs upon participating students compared with matched nonparticipating students.

In the current article, we report evidence from a national longitudinal study of training programs aimed at encouraging minority undergraduate students to pursue a research career in the biomedical sciences. Our focus is on a set of programs funded by the NIH under its Minority Biomedical Support Program (MBRS), Research Initiative for Science Excellence (RISE) mechanism. The MBRS program was founded in 1972 to provide research support for faculty and students at minority-serving educational institutions. RISE programs typically receive around $600,000 per year to support about 25 undergraduates and 5 master’s-level graduate students. Although each grant-receiving campus has flexibility in structuring its program, RISE programs typically include faculty mentoring of students, on-campus research opportunities, graduate school preparation, summer research internships, funding to attend and present at professional conferences, and substantial annual stipends. At the time of this writing, there were 40 existing RISE programs (35 in the United States, 5 international—Puerto Rico, the Virgin Islands, and Guam) in both public and private schools and at both 2-year and 4-year institutions.

Current Study

The purpose of the current study is to examine the effectiveness of a prototypical minority training program, the RISE program. We report initial results of *TheScienceStudy*, a prospective, national, longitudinal study of minority college students in the sciences. The data for the current article come from students enrolled at one of 25 four-year institutions with an NIH-funded RISE program. Study participants have been surveyed biannually (fall and spring semester) on issues pertaining to their interests and experiences, educational achievements, professional achievements, and career aspirations. *TheScienceStudy* began data collection in the fall semester of 2005 (Wave 0) and has continuously collected data through the present time (with funding through 2013). The initial panel of college students was recruited with the help of faculty and staff at our 25 partner campuses. At the time of recruitment, all panel members were attending a 4-year university, majoring in a biomedical discipline (e.g., biology, chemistry, psychology), and expressed interest in pursuing a science-related research career. A portion of the panel was recruited from minority training programs (primarily the NIH-funded RISE program, but some students were also supported by other federal, state, or privately funded programs). The remainder of the panel was recruited as a propensity score matched control sample. The matched sample participants were recruited from upper-division gateway science courses, such as organic chemistry.

In order to identify an appropriate matched sample, we utilized a propensity score matching procedure (Rosenbaum & Rubin, 1983, 1984; West et al., 2008). The purpose of a propensity score is to provide unbiased estimates of treatment effects in a quasi-experimental design. *TheScienceStudy* conducted a large-scale recruitment survey (prior to Wave 0) to identify a potential matched panel (N = 2,166). Propensity scores were created based on 11 variables (see Measure section for details). Although propensity scores are generally used to create a one-to-one match between treatment and control, in our analyses we use propensity scores as a covariate in order to remove any baseline correlation between the dependent variable (i.e., Intention) and treatment (i.e., RISE membership; Winship & Morgan, 1999). Propensity score is also used as a covariate out of practical considerations, as panel members

1. move in or out of minority training programs,
2. graduate from school at different rates, or
3. leave the panel.

The longitudinal design of *TheScienceStudy* allows us to address the issue of change over time among minority training program members. Furthermore, the propensity score matched sample allows us to compare the performance of RISE students against that of a comparison group. Our data analyses focused on three hypotheses. First, we hypothesized that the RISE and match groups would not differ on the Dependent Variable at Time 0 (no intercept differences between groups), after controlling for propensity to be in the RISE program. Second, we hypothesized that the RISE and match groups would differ in their growth trajectories over time, such that RISE students would show a higher level of intention to persist than the matched students. Finally, we hypothesized that two common elements of training programs, mentorship and research experience, would have unique and positive effects on sustaining student intentions over time.

Method

Participants

The data for the current study were collected in bi-annual (i.e., fall and spring semesters) surveys over a 3-year period, from fall 2005 (Wave 0) through fall 2007 (Wave 4). The analytic sample reported in this article consists of college students in their junior or senior year at Wave 0 (N = 801). The focus on upper-class undergraduates was determined by our focus on RISE programs, which typically recruit students in their junior or senior years (i.e., after students have picked a science-related major). The RISE students recruited into our study were, therefore, primarily in their junior or senior year. We chose to restrict the sample to those who started the survey during their junior or senior year in order to maximize similarity between the RISE and match groups. We also restricted the sample to students who were enrolled at a university with a RISE program (reduced sample to N = 647, in k schools = 24). We applied this restriction to ensure parity between campus-level affordances and resources available to the students. Finally, we restricted the sample to students who were either enrolled in a RISE program or were part of a matched sample (N = 469) at Wave 0. Students who were supported by another training program (e.g., privately funded programs, other NIH programs, NSF programs) were excluded from the current analyses. Finally, it is important to note that the data set was limited to responses provided by students during their undergraduate tenure (i.e., responses provided after graduation are not included). It is important to note that 58% of students provided data at three or more time points.

Throughout the 3 years reported here, some of the students moved in to (or out of) a funded training program. At initial enrollment, the sample consisted of a RISE group (n = 157) and matched group (n = 312) of participants. Throughout the course of data collection, some students in the matched group enrolled in a RISE program, and some students dropped out of the RISE program. For the purpose of clarity, our analyses were conducted on only those students who were (a) initially and continuously enrolled in the RISE program, (b) continuously not enrolled in any minority training program (i.e., the matched sample), or (c) initially in the match group but became enrolled in the RISE program during the survey. These restrictions resulted in a slightly smaller final sample of students: RISE n = 120 and match n = 295.

At enrollment (Wave 0), students in the RISE group were primarily in their early 20s (M = 22.10, SD = 3.69), female (66%), and split between those who self-identified as African American (48%), Hispanic/Latino/a (38%), Native American (1%), or "other" (2%). Similarly, students in the matched group were primarily in their early 20s (M = 22.28, SD = 3.15), female (69%), and split between African American (47%), Hispanic/Latino/a (40%), Native American (1%), or "other" (2%).

Measures

Our primary focus in this article is on student intentions to pursue a career as a research scientist. Although we recognize the limitations of focusing on intention, a sizable volume of psychological research has shown the empirical link between intention and behavior (Lent et al., 2005). In addition, focusing on intentions allows for statistical analyses that model growth over time and to explore the program-level activities that correlate with changes in intention.

Intention—Students were asked, "To what extent do you intend to pursue a science-related research career?" The response options ranged from 0 (*definitely will not*) to 10 (*definitely will*). For our longitudinal analyses below, this measure served as our outcome variable.

Propensity score—A propensity score variable was generated for all participants to control for baseline differences between students participating in the RISE program and students not participating in a training program.

We used baseline data from our samples to calculate propensity scores. We calculated a logistic regression model using 11 background variables and all two-way interactions to predict membership in the RISE program: age, gender, race/ethnicity, GPA, major, school, intention to pursue a scientific research career, educational progress (e.g., lower or upper division undergraduate, master's, or doctoral student), English as a first language, first generation attending college, and transfer status (from a community or junior college).[1] From the resulting logistic regression equation, we calculated the predicted probability of

[1]The variables that were uniquely statistically significant in the final model were GPA, age, transfer student, intention, gender, progress in school, school grouping, age squared, Gender × GPA, Gender × Transfer, and Gender × Progress in school.

membership in the RISE treatment group for each participant, which ranged from 0 to 1: RISE-treatment ($M_{propensity}$ = .47, *SD* = .18) and matched groups ($M_{propensity}$ = .41, *SD* = .15). The resulting propensity scores correctly classified 73% of the students in our panel. For the match students, the correct classification was 85%; for RISE students, the correct classification was 39%. These classification figures show the variability among students in both the match and RISE groups and the potential for overlap.

Minority training program status—At each wave of data collection, students were asked, "Are you currently enrolled in a minority training program at your college?" Responses at each wave were dummy-coded into the "match" variable: 0 = *RISE group* and 1 = *matched group*. Students who reported enrollment in the RISE program were coded as such, whereas those who reported no participation in any training program were assigned to the match group. Furthermore, students who reported support from another minority training program were excluded from the current analyses.

Scientific mentor—At each wave of data collection (except Wave 0), students were asked if there was a "faculty member," "program staff member," "graduate student," "postdoctoral fellow," or "scientific professional outside of the university" that they considered to be a mentor. Responses at each wave were dummy-coded into the "Science Mentor" variable: 0 = *no scientific mentor* and 1 = *has a scientific mentor*.

Research experience—At each wave of data collection (except Wave 0), students were asked if they had experience with "hands-on research activities with laboratory equipment in class," "worked in laboratory" at their current or other university, or had "worked on research at another location." Responses at each wave were dummy-coded as the "Research Experience" variable: 0 = *no research experience* and 1 = *research experience*.

Plan of Analysis

To test our hypotheses (i.e., difference between RISE and match groups, and effects of program elements on student intentions over time), we conducted a series of analyses in a hierarchical linear modeling framework (Raudenbush & Bryk, 2002). First, we used a model-building approach to identify the growth model that provides the best fit to the longitudinal data (Hox, 2002). Second, we tested the hypothesized equivalence of the RISE and match group intentions to pursue a scientific research career at Wave 0, controlling for propensity to be in a minority training program. Third, we tested the hypothesized difference between the RISE and match groups' growth trajectories over time, controlling for propensity score. Finally, we tested the persistent positive effects of mentorship and research experience on student intentions over time. All analyses were conducted using maximum likelihood estimation in Mixed-Models SPSS Version 17 (Statistical Package for the Social Sciences, 2008). Model fit was evaluated using the Akaike's information criterion (AIC; Akaike, 1974).

Variables

The following growth model analyses include six variables. The outcome variable is the student's intention "to pursue a science-related research career" score (labeled INTENTION in the models listed below). Five within-student (Level 1) predictors were entered into the model. The first growth variable (labeled Time.Linear) models the linear growth in student intention to pursue a science-related research career over five waves of data collection (Wave 0, 1, 2, 3, and 4). The linear growth variable was centered at initial enrollment in the study (i.e., Wave 0). The second growth variable (labeled Time.Quadratic) models the change in the growth trajectory. The third variable (labeled Match) was coded as a time-varying covariate to model the difference between the RISE and match groups (cf. McCoach

& Kaniskan, 2010). For example, students who were continuously part of the match group have a Match score profile as follows: 1, 1, 1, 1, 1. A student continuously enrolled in the RISE program would have a Match score profile of 0, 0, 0, 0, 0. Finally, a student who transitioned to (and stayed continuously enrolled in) the RISE group in the third wave of data collection would have a Match score profile of 0, 0, 1, 1, 1. The fourth variable (labeled Sci.Mentor) was coded as a time-varying covariate to model the persistent effect of mentorship on student intentions over time (McCoach & Kaniskan, 2010). For example, a student who reported having a mentor for the two consecutive semesters of his or her senior year would have a Sci.Ment score profile of 0, 0, 0, 1, 2. Since participants were not asked about mentors in the initial wave of data collection, scores in Wave 0 were set to 0. As with the science mentor variable, a fifth Level 1 variable (labeled Res.Exp) was coded as a time-varying covariate to model the persistent effect of engaging in research activities. Because participants were not asked about their research activities in the initial wave of data collection, scores at Wave 0 were set to 0. For example, a student who reported engaging in research activities for the three consecutive semesters prior to graduation would have a Res.Exp. score profile of 0, 0, 1, 2, 3.

One between-student (Level 2) predictor was entered into the model. The propensity score variable (centered at the grand mean) that expresses a student's likelihood of being enrolled in a RISE program (labeled PROPEN) was entered in the model to control for differences between RISE and matched students' intercept and growth trajectories.

Results

Descriptive Statistics

Prior to testing our hypotheses, we examined the patterns of change over time at the group level and at the individual level. As shown in Table 1, the sample sizes for the RISE and match groups fluctuated across waves, as some students graduated or transitioned from the match group into the RISE group. Furthermore, at Wave 0, the majority of students in both RISE (67%) and match (58%) groups rated their intention to pursue a science-related career as a 9 or 10 (scale range = 0 to 10; see Table 1). However, students in the RISE group consistently expressed high levels of intentions over time, whereas the proportion of match students expressing high intentions declined over time (see Table 1).

We also examined the group-level patterns of involvement in the two program elements of interest: scientific mentorship and research experience. As previously stated, these variables were coded to capture the persistent effect of mentorship and research on intentions. As shown in Table 1, the proportion of RISE students who report having a mentor or engaging in research activities increased steadily each semester. By the final wave of data collection, 94% of RISE students reported spending two or more semesters with a mentor, and 100% of RISE students reported engaging in two or more semesters of research. By contrast, students in the match group reported less systematic access to mentors and lower rates of engagement in research. By the final wave of data collection, 36% of match students reported spending two or more semesters with a mentor, and 41% reported engaging in two or more semesters of research activities.

Finally, we examined the individual-level patterns of change in student intentions to pursue a science-related research career. We compared each student's initially stated intention (Wave 0) to his or her last response as an undergraduate. We categorized students into three groups based on their intention scores: low (0–6), medium (7–8), and high (9–10). We found that individuals in the RISE program exhibited a high degree of stability (i.e., 86% of RISE students with high intentions at Wave 0 expressed high intentions in their last response) or positive growth (i.e., 52% of RISE students who expressed medium intentions at Wave 0 expressed high intentions in their last response). See Table 2. Conversely, students in the match group exhibited relatively less stability (i.e., 57% of match students with high intentions at Wave 0 expressed high intentions in their last response) and relatively less positive growth (i.e., 27% of match students who expressed medium intentions at Wave 0 expressed high intentions in their last response). See Table 2.

Together, these descriptive findings indicate that although the two groups are similar at the outset of the study, they diverge over time. Furthermore, this pattern shows that students in the RISE program are systematically connected with mentors and engaged in research activities, whereas their match counterparts have more sporadic access to both. We proceeded from these initial descriptive findings to formal statistical tests of our hypotheses.

Models

Statistical analyses were conducted within a multilevel modeling framework. The goals of the multilevel models presented below were to identify a Level 1 growth model that provides the best fit to student intention “to pursue a science-related research career” over time (Models 1–3), to compare the intercept and growth trajectories of RISE and matched students (Models 4 and 5), and to evaluate the effects of scientific mentorship (Model 6) and research experience (Model 7) in explaining the differences between the growth trajectories.

Modeling Growth Trajectories

Model 1a: The null model—The primary purpose of the null model (a.k.a. Random Effects ANOVA) is to estimate the intraclass correlation (ICC), which expresses both the proportion of variance that exists between students and the expected correlation between any two randomly chosen units (i.e., intention scores) within a cluster (i.e., student; Hox, 2002; Raudenbush & Bryk, 2002). Before proceeding, it will be advantageous to describe the multilevel model in terms of two sets of equations that specify predictions within students (Level 1) and between students (Level 2). The Level 1 null model is

$$\mathrm{INTENTION}_{ti}=\pi_{0i}+e_{ti},$$

where $\mathrm{INTENTION}_{ti}$ is the intention score for student i at time t, π_{0i} is intercept of the regression equation predicting intention score for student i (i.e., mean score across all time points), and e_{ti} is the deviation of student i at time t from his or her own average score across all time points. The Level 2 equation is

$$\pi_{0i}=\beta_{00}+r_{0i},$$

and the combined Level 1 and Level 2 equation is

$$\mathrm{INTENTION}_{ti}=\beta_{00}+r_{0i}+e_{ti},$$

where $\mathrm{INTENTION}_{ti}$ is the intention score for student i at time t, β_{00} is the intercept of the regression equation (i.e., mean score for all students at all time points), r_{0i} is the deviation of student i from the mean intention score of all students at all time points, and e_{ti} is the deviation of student i at time t from his or her own average score across all time points.

Examination of the fixed effect (β_{00} = 8.09, maximum is 10) indicates that the mean intention score across all students and time points was statistically significantly different from zero (see Table 3). More important, the ICC is calculated by partitioning the total variability of intention scores into two variance components: var(r_{0i}) = τ_{00} and var(e_{ti}) = σ^2. The estimate of between-student variance was τ_{00} = 2.47, and the estimate of within-student variance was σ^2 = 3.23. In our sample, the ICC is .43 (calculated as $\rho = \tau_{00} / (\tau_{00} + \sigma^2)$ $\rho=\tau_{00} / (\tau_{00} + \sigma^2)$ or 2.47/5.70 = 0.43), which indicates that almost half of the variability in scores lies between students. Furthermore, these findings indicate that the average correlation of scores within students is moderately high (.57). Therefore, we can expect that both intrapersonal factors (e.g., increasing engagement in research) as well as interpersonal factors (e.g., background characteristics, individual differences, or personality factors) may play important roles in explaining variability in student intentions.

Model 1b: The three-level null model—As described above, the implementation and specific features of the RISE programming can vary from campus to campus. To assess whether systematic variability in the scores was due to campus-level effects (k = 24 campuses), we ran a three-level null model to quantify the proportion of variability of intention scores within students (Level 1), between students (Level 2), and between campuses (Level 3). The three-level null model revealed that the estimate of within-student variance was 3.22, estimate of between-student variance was 2.33 (SE = .25, p < .001), and the estimate of between-campus variance was 0.16 (SE = .11, p = .08).[2] Furthermore, the Level 3 ICC was .02, indicating that only 2% of the variability in student intention scores is due to campus. Since campus effect on student scores is extremely small, we restrict the following analyses to a two-level model; however, future articles may examine campus-level effects more closely.

Model 2: Unconditional linear growth—The primary purpose of the unconditional linear growth model is to describe the linear change in student intention scores over time. The time variable (Time.Linear: centered at Wave 0) was added to the model in a two-phase process, wherein the linear growth slope (β_{10}) was estimated both with and without a random coefficient (r_{1i}). The random coefficient models the amount of variability across students around the average linear growth slope. We found that inclusion of the random coefficient improved the fit of the model (without r_{1i} AIC = 5,215.66; with r_{1i} AIC = 5,148.31), indicating that there is a statistically significant amount of variability across students around the mean linear growth slope. The Level 1 equation for the linear growth model is

$$\text{INTENTION}_{ti}=\pi_{0i}+\pi_{1i}(\text{Time.Linear})+e_{ti}.$$

The Level 2 models are

$$\pi_{0i}=\beta_{00}+r_{0i}\ \pi_{1i}=\beta_{10}+r_{1i},$$

and the combined model yields the following equation:

$$\text{INTENTION}_{ti}=\beta_{00}+\beta_{10}(\text{Time.Linear})+r_{0i}+r_{1i}(\text{Time.Linear})+e_{ti},$$

where INTENTION$_{ti}$ is the intention score for student i at time t (i.e., Wave 0), β_{00} is now the intercept of the regression equation predicting intention score for all students at Wave 0 (i.e., mean intention score at Wave 0), β_{10} is the linear effect of time on intention score or

[2]SPSS uses a two-tailed Wald statistic to test the variance components. The test of variance components should proceed using a one-tailed test. Therefore, we divided the p value provided by SPSS in half to derive the p value presented here.

mean linear growth slope, r_{0i} is now the deviation of student i from the intercept (i.e., mean across students), r_{1i} is now the deviation of student i from the mean linear growth slope, and e_{ti} is now the deviation of student i at time t from his or her own growth trajectory (Raudenbush & Bryk, 2002).

Examination of the fixed effects indicates that the intercept is still relatively high (β_{00} = 8.50) and that the linear growth slope is negative and statistically significant ($\beta_{10} = -0.31$) (see Table 3). Examination of the random effects indicates that the unconditional (i.e., without Level 2 predictors) linear growth model improved the fit of the model to the data. The estimate of within-person variance (σ^2) dropped from 3.23 (M1) to 2.54 (M2), indicating that linear growth explains 21% of the variance in intention scores within students over time (see Table 4). However, we were also interested in testing the curvilinear pattern of growth trajectories.

Model 3: Unconditional quadratic growth—The purpose of the quadratic growth model is to evaluate the curvilinear change in student intention scores over time. The quadratic time variable (Time.Quadratic) was estimated as a fixed effect because the models with a variance component would not converge. The Level 1 equation for the quadratic growth model is

$$\mathrm{INTENTION}_{ti}=\pi_{0i}+\pi_{1i}(\mathrm{Time.Linear})+\pi_{2i}(\mathrm{Time.Quadratic})+e_{ti}.$$

The Level 2 models are

$$\pi_{0i}=\beta_{00}+r_{0i}\ \pi_{1i}=\beta_{10}+r_{1i}\ \pi_{2i}=\beta_{20},$$

and the combined model yields the following equation:

$$\mathrm{INTENTION}_{ti}=\beta_{00}+\beta_{10}(\mathrm{Time.Linear})+\beta_{20}(\mathrm{Time.Quadratic})+r_{0i}+r_{1i}(\mathrm{Time.Linear})+e_{ti},$$

where β_{10} is now the instantaneous mean linear growth slope of all students at Wave 0 and r_{1i} is now the deviation of student i from the mean instantaneous linear slope, β_{20} is the mean quadratic growth slope of all students, and the interpretation of INTENTION$_{ti}$, β_{00}, r_{0i}, and e_{ti} remain the same as previous models.

Examination of the fixed effects (β_{00}, β_{10}, and β_{20}) shows that the parameter estimates were statistically significant (see Table 3). The intercept is relatively high (M3: β_{00} = 8.57), which indicates that initially, students expressed high levels of intentions to pursue a career in the sciences. The instantaneous linear growth slope is relatively large and negative (M3: $\beta_{10} = -0.57$), which indicates that student intentions dropped over time. The quadratic growth slope is relatively small and positive (M3: β_{20} = 0.08), which indicates that the decline of student intentions to pursue a scientific career leveled off or stabilized over time. Examination of the random effects indicates that the unconditional quadratic growth model improved the fit of the model to the data (see AIC in Table 4).

In summary, the unconditional growth models revealed that approximately half of the variability in student intentions to pursue a scientific career was attributable to between-subjects variance (ICC = .43). Furthermore, these models indicate that, on average, students initially expressed very high intentions to pursue a career in the sciences; however, their intentions declined relatively rapidly and eventually leveled off.

RISE and Match Growth Trajectories

Having ascertained our best fitting unconditional growth model, we began an examination of the impact of the RISE program on student intentions to pursue a scientific research career. As stated above, we did not expect to find differences between the RISE and match groups at the time of enrollment in the study, after controlling for the likelihood of being in a minority training program (PROPEN). We hypothesized differences in the growth trajectories of the RISE and match groups.

Model 4: RISE and match growth trajectories—The Match variable was sequentially entered at Level 1 as a predictor of the intercept, as interacting with the instantaneous linear growth slope, and as interacting with the quadratic growth slope. The Match variable provided the best fit to the data when predicting the intercept and interacting with the linear growth slope, but not the quadratic slope ($AIC_{Intercept}$ = 5,139.78 ; AIC_{Linear} = 5,137.14; $AIC_{Quadratic}$ = 5,138.90). The Level 1 equation is now

$$\mathrm{INTENTION}_{ti}=\pi_{0i}+\pi_{1i}(\mathrm{Time.Linear})+\pi_{2i}(\mathrm{Time.Quadratic})+\pi_{3i}(\mathrm{Match})+\pi_{4i}(\mathrm{Time.Linear}\times\mathrm{Match})+e_{ti}.$$

The Level 2 equations are

$$\pi_{0i}=\beta_{00}+r_{0i}\ \ \pi_{1i}=\beta_{10}+r_{1i}\ \ \pi_{2i}=\beta_{20}\ \ \pi_{3i}=\beta_{30}\ \ \pi_{4i}=\beta_{40},$$

and the combined model yields the following equation:

$$\begin{aligned}\mathrm{INTENTION}_{ti}=&\beta_{00}+\beta_{10}(\mathrm{Time.Linear})\\ &+\beta_{20}(\mathrm{Time.Quadratic})\\ &+\beta_{30}(\mathrm{Match})\\ &+\beta_{40}(\mathrm{Time.Linear}\\ &\times\mathrm{Match})+r_{0i}+r_{1i}(\mathrm{Time.Linear})+e_{ti},\end{aligned}$$

where β_{00} is now interpreted as the mean RISE student intention score at Wave 0, β_{30} is the difference between the RISE and Match mean intention score at Wave 0 (i.e., $\beta_{00} + \beta_{30}$ = Match group Intercept), β_{10} is now the instantaneous linear growth slope for the RISE group, β_{40} is now the difference between RISE and match groups' instantaneous linear growth slope (i.e., $\beta_{10} + \beta_{40}$ = match group instantaneous linear growth slope), and the interpretations of the β_{20}, r_{0i}, r_{1i}, and e_{ti} coefficients remain unchanged.

Examination of the fixed effects showed a marginally significant difference in initial intention scores of the RISE and match ($\beta_{30} = -0.40$, $SE = .21$, $p = .06$) groups. Although the difference between the groups only trended toward statistical significance, we proceeded with entering student propensity scores into the model to control for these initial differences.

Model 5: Controlling for propensity to be in a RISE—The propensity score variable (centered at the grand mean) was sequentially entered into the model as a Level 2 predictor of the intercept and growth slopes. Propensity score improved model fit when predicting the intercept but did not improve model fit when predicting the linear or quadratic growth slopes ($AIC_{Intercept}$ = 5,122.19, AIC_{Linear} = 5,123.40, and $AIC_{Quadratic}$ = 5,125.11). The Level 1 equation remains unchanged; however, the Level 2 equations are

$$\pi_{0i}=\beta_{00}+\beta_{00}(\text{PROPEN}_{\text{CGM}})+r_{0i}\ \pi_{1i}=\beta_{10}+r_{1i}\ \pi_{2i}=\beta_{20}\ \pi_{3i}=\beta_{30}\ \pi_{4i}=\beta_{40},$$

and the combined equation is

$$\begin{aligned}\text{INTENTION}_{ti}=&\beta_{00}+\beta_{01}(\text{PROPEN}_{\text{CGM}})\\&+\beta_{10}(\text{Time.Linear})\\&+\beta_{20}(\text{Time.Quadratic})+\beta_{30}(\text{Match})\\&+\beta_{40}(\text{Time.Linear}\\&\times\text{Match})+r_{0i}+r_{1i}(\text{Time.Linear})+e_{ti},\end{aligned}$$

where β_{00} is now interpreted as the mean RISE student intention score at Wave 0, controlling for propensity score; β_{30} is now the difference between the RISE and match mean intention scores at Wave 0, controlling for propensity score; β_{01} is the effect of propensity score on student intention at Wave 0; and the interpretation of the β_{10}, β_{20}, β_{30}, β_{40}, r_{0i}, r_{1i}, and e_{ti} coefficients remains unchanged.

Examination of the fixed effects indicates that the difference between RISE and match group intercepts is not statistically significant (M5: β_{30}), controlling for propensity to be in a minority training program (see Table 3). The instantaneous linear growth slopes for the both the RISE and match groups are still negative; however, the linear slope for the match group is still significantly more negative than the RISE group's linear slope (M5: β_{40}). This finding indicates that students in the match group exhibit a steeper decline in their intentions to pursue a scientific career compared to students in the RISE program. The quadratic growth slope remains unchanged, indicating that the decline in student intentions levels off over time regardless of program status. Examination of the random effects indicates that the conditional growth models improved overall fit to the data. The estimate of between-person variance around the intercept dropped by 10% after the inclusion of the predictors (τ_{00}: M3 = 1.68, and M5 = 1.51; see Table 4).

To provide a visual representation of the accuracy of our growth model for RISE and match students, we plotted the model predicted values and raw mean scores (see Figure 1). The visual plots show that our predicted values from hierarchical linear modeling (HLM) closely map on to the observed changes in student intention scores over time.

In summary, the first key finding from these models is that there are no differences between the RISE and match groups' intentions to pursue a scientific career at Wave 0, as hypothesized. In other words, the two groups appear to be relatively evenly matched; thus, we can more confidently make comparisons between these groups, particularly after covarying out student propensity scores. The second key finding is that the linear growth slopes of the RISE and match groups are different, such that the match group exhibits a steeper decline in their intentions to pursue a scientific research career over time (Hypothesis 2 supported). As shown in Figure 1, the Wave 0 predicted intention score for RISE students is 8.71 and is 8.52 for match students; however, by Wave 4, the predicted intention scores are 8.37 and 7.20, respectively. The model indicates that on average, RISE program members express a 0.38 point drop in their intention to pursue a scientific career, while on average, match group members express a 1.22 point drop in their intention to pursue a science-related research career.

Model 6: Effect of scientific mentorship—Next, we examined the effects of two common features of minority training programs. The scientific mentorship variable was sequentially entered as a Level 1 time-varying covariate; however, it was not a statistically significant predictor (Estimate = 0.05, *SE* = .12, *p* = .69) and did not improve model fit (AIC = 5,124.03). Having a scientific mentor did not produce a unique effect on student intentions over time; therefore, the mentorship variable was dropped from the model.

Model 7: Effect of research experience—The research experience variable was entered as a Level 1 time-varying covariate. Research experience improved model fit (AIC = 5,113.03). The Level 1 equation is now

$$\begin{aligned}\text{INTENTION}_{ti}=&\pi_{0i}+\pi_{1i}(\text{Time.Linear})\\&+\pi_{2i}(\text{Time.Quadratic})\\&+\pi_{3i}(\text{Match})\\&+\pi_{4i}(\text{Time.Linear}\\&\quad\times\text{Match})\pi_{1i}(\text{Res.Exp})+e_{ti}.\end{aligned}$$

The Level 2 equations are

$$\pi_{0i}=\beta_{00}+\beta_{01}(\text{PROPEN}_{\text{CGM}})+r_{0i}\ \pi_{1i}=\beta_{10}+r_{1i}\ \pi_{2i}=\beta_{20}\ \pi_{3i}=\beta_{30}\ \pi_{4i}=\beta_{40}\ \pi_{5i}=\beta_{50},$$

and the combined equation is

$$\begin{aligned}\text{INTENTION}_{ti}=&\beta_{00}+\beta_{01}(\text{PROPEN}_{\text{CGM}})\\&+\beta_{10}(\text{Time.Linear})\\&+\beta_{20}(\text{Time.Quadratic})\\&+\beta_{30}(\text{Match})\\&+\beta_{40}(\text{Time.Linear}\\&\quad\times\text{Match})+\beta_{50}(\text{Res.Exp})\\&+r_{0i}+r_{1i}(\text{Time.Linear})+e_{ti},\end{aligned}$$

where β_{50} is the persistent effect of research experience on intentions over time, and the interpretations of β_{00}, β_{01}, β_{20}, β_{30}, β_{40}, r_{0i}, r_{1i}, and e_{ti} remain essentially unchanged, controlling for research experience.

An examination of the fixed effects shows that research experience has a statistically significant positive effect on intention scores over time (M7: β_{50} = .51, p < .001), such that each additional semester of research experience dramatically attenuates the decline of student intentions (see Table 1). Interestingly, the effect is consistent for both RISE and match students, such that the buffering effect is equally strong for students supported by the RISE program and for nonsupported students. This effect is shown in Figure 2. Furthermore, after controlling for research experience, there is no longer a statistically significant difference between the linear slopes of the RISE and match students (see Table 1). In order to visually represent the persistent impact of research experience on student intentions, we plotted the model-predicted values for RISE and match students that accumulated three semesters of research experience or no semesters of research experience over the 3-year period in Figure 2. The visual plot clearly shows that RISE and match students with high levels of research experience sustain high intentions to pursue scientific research careers, while those who do not exhibit a substantial drop in their intentions.

In summary, we hypothesized that two common features of minority training programs—mentorship and research experience—would have uniquely positive effects on student intentions to pursue science-related research careers over time. We found that only research experience uniquely and strongly influenced the growth trajectories of students' intention to pursue a science-related research career.

Discussion

The results reported in this article represent the first attempt to evaluate the efficacy of federally funded minority training programs using a multisite, longitudinal, quasi-experimental design. Our focus was on the RISE program, and we provide longitudinal analyses of students from 25 funded programs. The results provide strong evidence for the ability of the RISE program to sustain student intentions to pursue a research career. Although the general slope in student intentions to pursue a research career is negative, students supported by the RISE program showed a higher level of persistence in their intentions over time and level off at a higher point than do students from the matched sample.

In our efforts to fit a growth model to student intentions, a very clear finding emerged: The trajectory was negative. At the point when the students joined the study, they all expressed some interest in pursuing a scientific research career (mean intention = 8.50, out of 10). However, the linear growth trajectory was −.31, indicating a steady decline in intentions over time. Because the coefficient is in unstandardized units, it can be directly interpreted such that each semester, the average student declines by .31 units in their intention (on the 10-point scale). Using both this linear and the quadratic growth trajectories, by Wave 4 (nearly 3 years after the initial recruitment), the modeled student intention has dropped to 6.61. This finding is consistent with prior studies showing that although many university students express an interest in science, these interests decline steadily during their years as an undergraduate (Hueftle et al., 1983; Hurtado et al., 2008).

Our results also show the strong potential for a science training program to moderate the decline in student intentions to pursue a research career. Using our propensity score matching procedure, we compared students who were supported by a RISE program with similar students who were not supported by any program. Although both RISE and match students showed a negative growth trajectory, the RISE students declined markedly less. Importantly, these results speak to the "buffering" effect of the training program, rather than a "whetting of interests." That is, students were already interested in science at the initial wave, and their participation in the RISE program served to sustain this interest over time.

Our results show clear evidence that the RISE program can sustain student interest in the sciences. But what is it about the RISE program that produces these effects? What is the generative mechanism? Like most science training programs, campuses are given latitude in designing and structuring program activities, and there was considerable variation in programming features across our 25 campuses. As a starting point, we began by examining two common program elements: research experience and mentorship. For research experience, our analyses clearly show that students who report participating in research as undergraduates are substantially more likely to sustain their interests in science. Importantly, the effect of research experience was not limited to students from a funded program. Although students from RISE programs were substantially more likely to show a sustained interest in science, nonfunded students who engaged in research also showed more interest in science than did matched students without this experience. It is important to point out that given the nature of our design and analyses, our results do not show that research experience promotes an intention to pursue a research career. But rather, for those students who already have the intention, research experience can help to sustain it.

These results echo a growing body of evidence showing the impact of research experiences on young science students (Morley, Havick, & May, 1998; Nagda, Gregerman, Jonides, von Hippel, & Lerner, 1998; NSF, 1989; Russell, Hancock, & McCullough, 2007; Seymour, Hunter, Laursen, & DeAntoni, 2004). However, existing data about the impact of research experience on undergraduates are based on small samples or rely on retrospective accounts (Hackett, Croissant, & Schneider, 1992; Zydney, Bennett, Shahid, & Bauer, 2002). Our longitudinal results add to these prior studies in showing that undergraduates who participate in research have a sustained intention to pursue a research career. Unlike prior studies that ask students to reflect back on their research experiences or to rate the immediate impact of a special summer or research-intensive program, our results show that participating in research exerts a direct impact on student intentions and that this effect is durable across time. Furthermore, our results clearly show that continued engagement in research activities has a persistent and additive effect on student intentions over time (i.e., more is better).

An important question that remains to be addressed is *why* research experiences increase academic persistence? On one hand, there is some evidence to suggest that research experience results in higher student self-efficacy. Using instruments, collecting data, adhering to research protocols, and hands-on experience with the process of science serve to increase a student's perception of his or her ability as a scientist (Chemers, Hu, & Garcia, 2001; Lent et al., 2005). However, recent analyses using data from *TheScienceStudy* suggest that self-efficacy is just a small piece of the persistence effect. Estrada-Hollenbeck et al. (2010) suggest that a student's identity as a scientist and the degree to which he or she values the objectives of science serve as a stronger explanatory variable for the persistence effect. Similarly, analyses by Merolla et al. (under review) show that joining a minority training program precedes a change in student identity as a scientist. That is, as a result of joining a science training program, students are more likely to think of themselves as scientists. And it is this change in identity that sustains their interest in pursuing a scientific research career.

Our analyses also examined the impact of having a mentor on student intentions, and surprisingly, our results suggest that having a mentor does not affect student intentions. Prior studies have generated mixed results with regard to the impact of mentorship on student interest and academic performance. Although some studies have touted the benefits of having a mentor (Chickering & Gamson, 1991; Fries-Britt & Turner, 2002; Kuh & Hu, 2001), several studies have shown null effects (Haring, 1999). Our results suggest that just *having* a mentor is not sufficient. Rather, it is likely that some mentoring relationships are better than others (Algería, 2009; Denofrio, Russell, & Lopatto, 2007; Pfund, Pribbenow, Branchaw, Lauffer, & Handelsman, 2006).

Caveats and Limitations

Although it is tempting to interpret our results as broad-level support for the efficacy of minority training programs, we want to caution against generalizing too far from our data. Our study focused on the RISE program—a long-standing, federally funded program aimed at promoting diversity among the scientific research community. Our findings provide evidence for the efficacy of the program, but from the results reported above we cannot pinpoint all of the program elements that are linked with success. Although there are hundreds of programs nationally with a similar goal of promoting diversity (e.g., NSF-funded, state-funded, institution-funded, and those funded locally through foundations and

NIH-PA Author Manuscript

NIH-PA Author Manuscript

private donations), they differ tremendously in their components. Clearly some of these programs will be more effective than others, and additional data are needed to identify the mediating mechanisms that explain program-level success.

Although there are many strengths to the design of this longitudinal prospective study, there are also limitations. One such limitation is that the causal ordering between research experience and intentions is uncertain since students are not randomly assigned to research conditions. Furthermore, the hierarchical linear modeling framework is not well suited to disentangle the potential feedback loop between research and intentions. However, our data are consistent with a causal interpretation that engagement in research would affect intentions to pursue science careers over time.

A second, potentially more serious limitation of the current study concerns the sample characteristics. Our analytic selection criteria limited responses to RISE and match students only during their undergraduate years. This restriction reduced the wave-to-wave sample size of the RISE group to a fairly small number of students. This limitation might indicate that the findings do not generalize to the wider population of RISE students. However, we would counter this critique with the fact that 61% of our RISE sample (and 45% of the match sample) had graduated with a baccalaureate degree by Wave 4. Therefore, the current analyses do not reflect attrition from the sample but, rather, the expected transition out of the undergraduate educational pipeline.

A third limitation to the current study is that our coding scheme for research activities does not specify the varied types of research activities that can affect student intentions. It is possible that certain types of research activities are more impactful compared to others. Future studies should look more closely into the types of research activities that are most impactful (e.g., faculty- vs. student-generated research projects).

A fourth limitation to the current study was the role of higher level contextual factors, such as program-specific effects and institutional effects. It is plausible that some RISE programs or some institutions exert differential influence on their students' intentions to pursue a scientific research career. Although our current study did not find empirical evidence of a program/institution-level effect on student intentions (i.e., "*Model 1b: The Three-Level Null Model*" did not exhibit a statistically significant variance component at the institution level), we believe this is an area that needs more study.

In closing, there have been ongoing efforts to promote diversity among the scientific research community for more than 40 years. However, little is known about the efficacy of these programs, and at a national level, we continue to see sizable disparities in the educational achievements of students from different racial and ethnic groups. Education researchers have been involved in these efforts since the beginning, but to date, we have not applied the rigorous standards of education research and quantitative methodology to evaluate and inform these programs. We believe that this is a place where policymakers are eager to hear from social and behavioral scientists, and we encourage our research community to embrace these questions.

Acknowledgments

Our appreciation goes to Victor Rocha and the staff at the Office for Biomedical Research and Training (OBRT) at CSUSM for their ongoing support. We also want to acknowledge the important contributions of Brian McDonald and the Survey Research Lab at Kent State University; also to our collaborators at Kent State University, David Merolla, and Steve Mullet; and finally, to our student research assistants at CSUSM: Sara Aguilar, Heather Butler, Nancy Larson, and Adam Zaleski. Portions of this article were presented at the 2008 and 2009 meetings of the Association for Psychological Science, in Chicago and San Francisco, respectively.

Financial Disclosure/Funding

The authors disclosed receipt of the following financial support for the research and/or authorship of this article: Funding for this study was provided by a grant from the National Institute of General Medical Sciences (R01-GM075316) to the first author.

References

Akaike H. A new look at the statistical model identification. IEEE Transactions on Automatic Control. 1974; 19:716–723.

Algería M. Training for research in mental health and HIV/AIDS among racial and ethnic minority populations: Meeting the needs of new investigators. American Journal of Public Health. 2009; 99:26–30.

Bickman, L., editor. Validity and social experimentation: Donald Campbell's legacy. Thousand Oaks, CA: Sage; 2000.

Cabrera AF, Nora A, Castaneda MB. College persistence: Structural equation modeling test of an integrated model of student retention. Journal of Research in Higher Education. 1993; 64:123–139.

Campbell, DT. The experimenting society. In: Campbell, DT.; Overman, S., editors. Methodology and epistemology for social science: Selected papers. Chicago: University of Chicago Press; 1988. p. 290-314.

Chemers M, Hu L, Garcia B. Academic self-efficacy and first-year college student performance and adjustment. Journal of Educational Psychology. 2001; 93:55–64.

Chickering, AW.; Gamson, ZF. Applying the seven principles for good practice in undergraduate education. San Francisco, CA: Jossey-Bass; 1991.

Collea FP. Increasing minorities in science and engineering: A critical look at two programs. Journal of College and Science Teaching. 1990; 20:31–41.

Cook, BJ.; Córdova, DI. Minorities in higher education: Twenty-second annual report. Washington, DC: American Council on Education; 2006.

Denofrio L, Russell B, Lopatto D. Mentoring: Linking student interests to science curricula. Science. 2007; 318:1872–1873. [PubMed: 18096791]

DePass, A.; Chubin, D. Understanding interventions that encourage minorities to pursue research careers. Bethesda, MD: American Council for Cell Biology; 2008.

Estrada-Hollenbeck M, Woodcock A, Hernandez P, Schultz PW. Toward a social influence framework that explains minority student integration into the scientific community. Journal of Educational Psychology. 2010

Fries-Britt S, Turner B. Uneven stories: Successful black collegians at a black and a white campus. The Review of Higher Education. 2002; 25(3):315–330.

Hackett EJ, Croissant J, Schneider B. Industry, academe, and the values of undergraduate engineers. Research in Higher Education. 1992; 33:275–295.

Hanson, S. Swimming against the tide: African American girls and science education. Philadelphia, PA: Temple University Press; 2009.

Haring MJ. The case for a conceptual base for minority mentoring programs. Peabody Journal of Education. 1999; 74(2):5–14.

Harrell PM, Forney WS. Ready or not, here we come: Retaining Hispanic and first-generation students in postsecondary education. Community College Journal of Research and Practice. 2003; 27:147–156.

Hausmann LRM, Schofield JW, Woods RL. Sense of belonging as a predictor of intentions to persist among African American and White first-year college students. Research in Higher Education. 2007; 48:803–845.

Hox, J. Multilevel analysis: Techniques and applications. Mahwah, NJ: Erlbaum; 2002.

Hueftle, JJ.; Rakow, SJ.; Welch, WW. Images of science. Minneapolis: University of Minnesota Science Assessment Research Project; 1983.

Hurtado S, Eagan M, Cabrera N, Lin M, Park J, Lopez M. Training future scientists: Predicting first-year minority student participation in health science research. Research in Higher Education. 2008; 49:126–152. [PubMed: 23503996]

Johnson, D., editor. Minorities and girls in school: Effects on achievement and performance. Thousand Oaks, CA: Sage; 1997.

Jun A, Tierney WG. At-risk urban students and college success: A framework for effective preparation. Metropolitan Universities. 1999; 9:49–60.

Kuh GD, Hu S. The effects of student-faculty interaction in the 1990s. Review of Higher Education. 2001; 24(3):309–332.

Lam PC, Mawasha PR, Doverspike D, McClain B, Vesalo J. A description and evaluation of the effects of a preengineering program for underrepresented, low-income, and/or first generation college students and the University of Akron. Journal of Women and Minorities in Science and Engineering. 2000; 6:221–228.

Lent RW, Brown SD, Sheu HB, Schmidt J, Brenner BR, Gloster CS, Treistman D. Social cognitive predictors of academic interest and goals in engineering: Utility for women and students at historically black universities. Journal of Counseling Psychology. 2005; 52:84–92.

Maton KI, Hrabowski FA III. Increasing the number of African American Ph.D.s in the sciences and engineering: A strengths-based approach. American Psychologist. 2004; 59:629–654.

Maton, KI.; Hrabowski, F.; Ozlemir, M.; Wimms, H. Enhancing representation, retention, and achievement of minority students in higher education: A social transformation theory of change. In: Shinn, M.; Yoshikawa, H., editors. Toward positive youth development: Transforming schools and community programs. New York, NY: Oxford University Press; 2008. p. 115-132.

McCoach DB, Kaniskan B. Using time-varying covariates in multilevel growth models. Frontiers in Psychology. 2010; 1:1–12. [PubMed: 21833184]

Merolla, D.; Serpe, R.; Stryker, S.; Schultz, PW. Structural precursors to identity processes: The role of proximate social structures. Kent State University; (under review). Unpublished manuscript.

Mervis J. NIH program gives minorities a chance to make their MARC. Science. 2003; 301:455–456. [PubMed: 12881544]

Mervis J. Biomedical training programs: NIH told to get serious about giving minorities a hand. Science. 2006; 311:328–329. [PubMed: 16424312]

Morley RL, Havick JJ, May GS. An evaluation of the Georgia Tech summer undergraduate program of research in electrical engineering for minorities. Journal of Engineering Education. 1998; 87:321–325.

Nagda BA, Gregerman SR, Jonides J, von Hippel W, Lerner JS. Undergraduate student-faculty partnerships affect student retention. The Review of Higher Education. 1998; 22:55–72.

National Center for Education Statistics. Digest of education statistics. Washington, DC: U.S. Department of Education; 2005. Available at http://nces.ed.gov/programs/digest/d05/tables/dt05_009.asp

National Institutes of Health. Assessment of NIH minority research training programs: Phase 3. Washington, DC: National Academies Press; 2005. Available at http://www.nap.edu/catalog.php?record_id=11329

National Science Foundation. Report on the National Science Foundation disciplinary workshops on undergraduate education. Washington, DC: Author; 1989.

National Science Foundation. Women, minorities, and persons with disabilities in science and engineering: 2000 (Publication No. NSF-00–327. Arlington. VA: Author; 2000. Retrieved from ERIC database (ED 449025)

Orfield, G.; Losen, D.; Wald, J.; Swanson, C. Losing our future: How minority youth are being left behind by the graduation rate crisis. Cambridge, MA: The Civil Rights Project at Harvard University; 2004.

Pfund C, Pribbenow CM, Branchaw J, Lauffer SM, Handelsman J. The merits of training mentors. Science. 2006; 311:473–474. [PubMed: 16439648]

Raudenbush, SW.; Bryk, AS. Hierarchical Linear Models: Applications and data analysis methods. 2nd ed. Newbury Park, CA: Sage; 2002.

Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika. 1983; 70:41–55.

Rosenbaum PR, Rubin DB. Reducing bias in observational studies using subclassification on the propensity score. Journal of the American Statistical Association. 1984; 79(387):516–524.

Russell SH, Hancock MP, McCullough J. Benefits of undergraduate research experiences. Science. 2007; 316:548–549. [PubMed: 17463273]

Seymour E, Hunter AB, Laursen SL, DeAntoni T. Establishing the benefits of research experiences for undergraduates in the sciences: First findings from a three-year study. Science Education. 2004; 88(4):493–534.

Statistical Package for the Social Sciences. SPSS for Windows, Rel. 17.0.0 [computer software]. Chicago, IL: SPSS Inc; 2008.

U.S. General Accounting Office. Federal science, technology, engineering, and mathematics programs and related trends. 2005. Retrieved from http://www.gao.gov/new.items/d06114.pdf

West S, Duan N, Peguegnat W, Gaist P, Jarlais DD, Holtgrave D, Mullen PD. Alternatives to the randomized controlled trail. American Journal of Public Health. 2008; 98(8):1359–1366. [PubMed: 18556609]

Winship C, Morgan S. The estimation of causal effects from observational data. Annual Review of Sociology. 1999; 25:659–707.

Zydney A, Bennett J, Shahid A, Bauer K. Impact of undergraduate research experience in engineering. Journal of Engineering Education. 2002; 97:151–157.

Biographies

P. WESLEY SCHULTZ is Professor of Psychology at California State University, San Marcos. His research interests are in applied social psychology, and in using psychological theories and methods to understand and solve social issues. Recent projects have focused on science training programs, science education related to climate change, and behavioral solutions to environmental problems. wschultz@csusm.edu. Website: www.csusm.edu/schultz.

PAUL R. HERNANDEZ is a doctoral candidate in Educational Psychology at the University of Connecticut. His research interests include longitudinal design, latent variable statistical methods, psychometric validation, and applications of psychological theory to understand and address social problems. Recent projects have focused on science training programs, SEM methods for conducting multivariate meta-analysis, and the psychometric validation of measurements instruments for use with diverse populations.

ANNA WOODCOCK is a doctoral candidate in the Department of Psychological Sciences at Purdue University. Her research interests lie in the broad areas of diversity, prejudice and stereotyping. Specifically she is investigating: the impact of implicit racial and gender bias on behavior and strategies to reduce bias, the processes by which stereotype threat operates, and the psychological processes underlying the underrepresentation of women and minorities in STEM careers.

MICA ESTRADA earned her doctorate in Social Psychology from Harvard University (1997) and is currently a Senior Research Scientist at California State University, San Marcos. Her area of expertise is social influence, which includes the study of identity, values, and intergroup relations. Her recent research has focused on academic perseverance, climate change education, and identity alignment. mestrada@csusm.edu. Website: http://faculty.csusm.edu/mestrada.

RANDIE C. CHANCE is a doctoral student in the Department of Psychology at Southern Illinois University Carbondale. Her research interests include self/identity, specifically in

multicultural and student identity. Her recent research has focused on academic identity, academic persistence, and identity alignment. She completed her research for this article while at California State University, San Marcos. rchance@siu.edu.

MARIA AGUILAR is a graduate student at California State University, San Marcos. Her research interests include understanding effective interventions to address educational disparities among minorities, as well as the social psychological processes that underlie differences in educational attainment. maguilar@csusm.edu.

RICHARD T. SERPE is Professor and Chair of Sociology at Kent State University. His primary research continues to focus on the relationship between and identity/self and social action. He is currently engaged in a multi-year panel study funded by the National Institutes of Health assessing identity theory to changes in self-evaluation, the link between intention and behavior, and the interrelations of structural identity and social identity. He is also working on national survey funded by the National Science Foundation to investigate the differential relationship between, structural identity, identity control and self verification as predictors of anxiety and depression.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

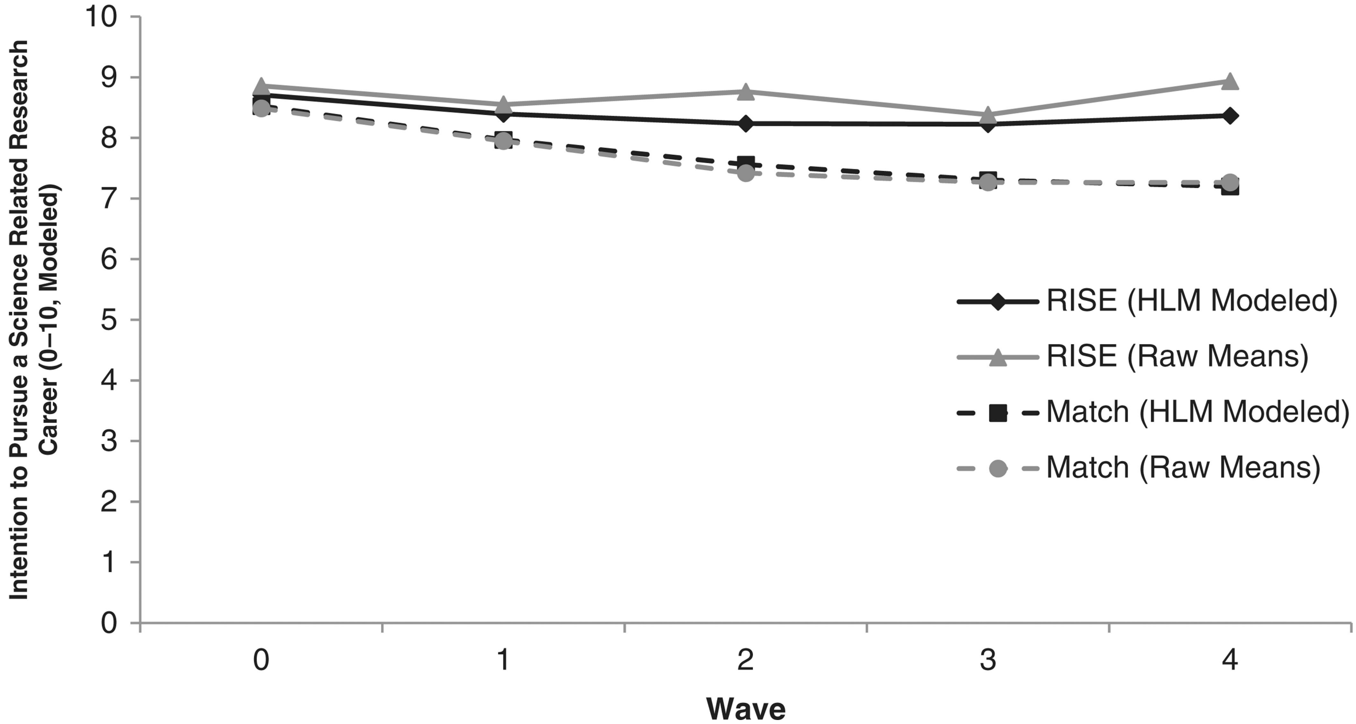

FIGURE 1.
Comparison of Raw Mean Scores and Growth Model Predicted Values of Student's Intentions "to Pursue a Science-Related Research Career" (M5: + Propensity Score).

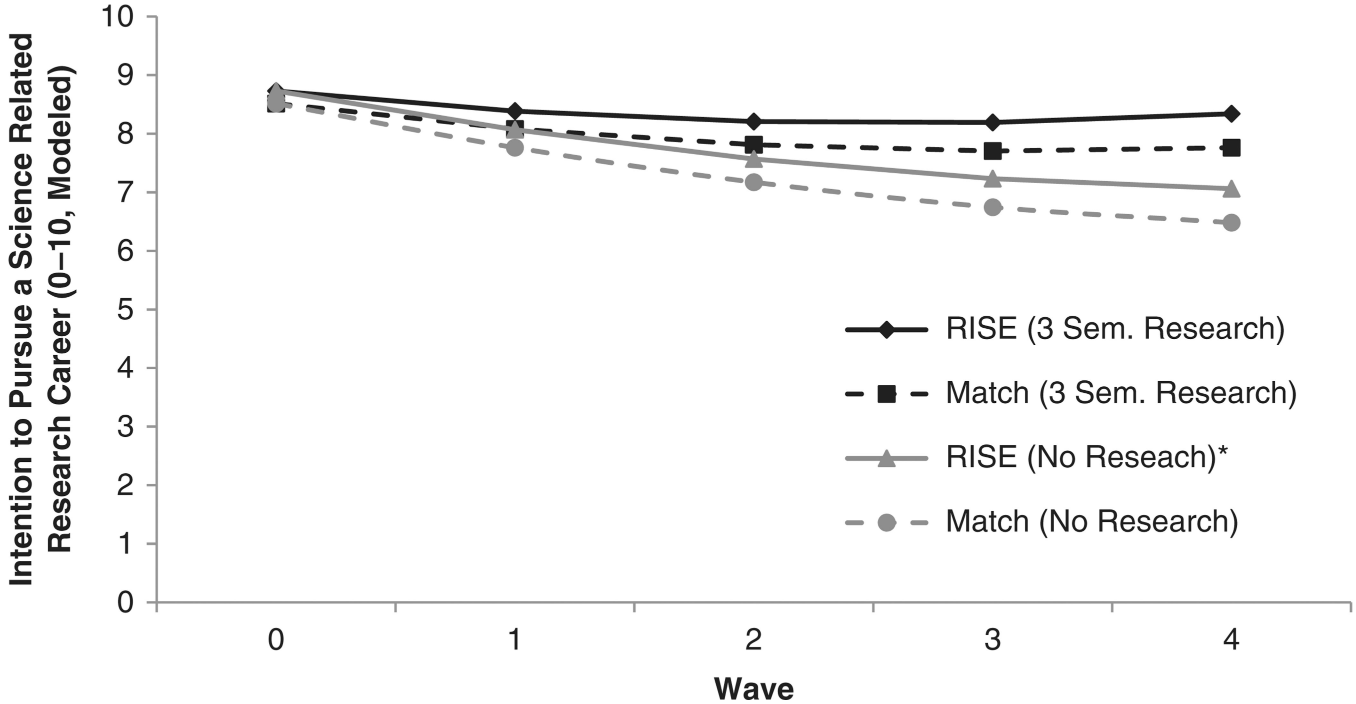

FIGURE 2.
The Effect of Research Experience on Intentions for Both RISE and Match Students, Controlling for Propensity Score (M6: + Research Experience).
Note. Figure plots the cumulative effect of research on RISE and match student intention over time. RISE (with three semesters of research) and match (with three semesters of research) are the predicted profile over the 3-year period; match (No Research) = modeled profile for match students with no research expeience over the 3-year period; RISE (No Research)* = shows a hypothetical modeled profile of RISE students with no research experience (i.e., all students in the RISE program reported two or more semesters of research experience over the 3-year period).

NIH-PA Author Manuscript
NIH-PA Author Manuscript

TABLE 1

Descriptive Statistics of Students' Intentions "to Pursue a Science-Related Research Career," Cumulative Science Mentorship, and Cumulative Research Experience for RISE and Match Groups

		Wave									
Variable	**Descriptives**	**0**	**1**	**2**	**3**	**4**	**0**	**1**	**2**	**3**	**4**
Intentions				*RISE*					*Match*		
	N	104	60	63	63	31	310	171	167	150	78
	M	8.86	8.55	8.76	8.38	8.94	8.49	7.95	7.42	7.27	7.27
	SD	1.61	2.33	1.63	2.37	1.90	1.81	2.47	2.81	2.86	2.86
f	0–2	1%	3%	0%	3%	0%	1%	5%	9%	10%	9%
	3–4	1%	3%	3%	5%	3%	3%	8%	9%	5%	13%
	5–6	7%	5%	8%	13%	13%	10%	12%	12%	18%	21%
	7–8	24%	23%	22%	10%	6%	28%	22%	22%	24%	15%
	9–10	67%	65%	67%	70%	77%	58%	54%	48%	43%	42%
Cumulative Science Mentorship											
	M	0.00	0.92	1.43	2.21	2.84	0.00	0.61	0.90	1.16	1.37
	SD	0.00	0.28	0.59	0.79	0.93	0.00	0.49	0.77	1.18	1.30
f	0	100%	8%	5%	3%	3%	100%	48%	50%	43%	31%
	1		92%	48%	13%	3%		52%	25%	18%	33%
	2			48%	44%	23%			25%	20%	13%
	3				40%	48%				19%	14%
	4					23%					9%
Cumulative Research Experience											
	M	0.00	0.95	1.57	2.44	3.03	0.00	0.61	0.90	1.26	1.38
	SD	0.00	0.22	0.50	0.62	0.66	0.00	0.49	0.77	1.04	1.21
f	0	100%	5%	0%	0%	0%	100%	39%	35%	31%	28%
	1		95%	43%	6%	0%		61%	40%	26%	31%
	2			57%	43%	19%			25%	30%	22%
	3				51%	58%				13%	13%
	4					23%					6%

Note. N = number of participants; f indicates the proportion of responses at each value (or range of values) in the frequency distribution. Since no science mentorship or research activity information was collected in the initial wave of data collection (Wave 0), the values for Cumulative Science Mentorship and Cumulative Research Experience have been set to 0.

TABLE 2

Cross-Tabulations Showing Patterns of Individual Changes in Intention "to Pursue a Science-Related Research Career" for RISE and Match Undergraduates

		Last response							
		RISE (N = 104)				*Match (N = 244)*			
		Low	**Medium**	**High**	**Total**	**Low**	**Medium**	**High**	**Total**
First Response	Low (0–6)	9 (82%)	1 (9%)	1 (9%)	11 (100%)	22 (67%)	6 (18%)	5 (15%)	33 (100%)
	Medium (7–8)	6 (26%)	5 (22%)	12 (52%)	23 (100%)	35 (47%)	20 (27%)	20 (27%)	75 (100%)
	High (9–10)	5 (7%)	5 (7%)	60 (86%)	70 (100%)	31 (23%)	27 (2%)	78 (57%)	136 (100%)

Note. Sixty-six participants provided data at only Wave 0; thus, they were excluded from this analysis. First response scores were taken from Wave 0; last response scores were taken from the last data point at which the participants gave a response as undergraduates. Scores are split into low (range: 0–6), medium (range: 7–8), and high (range: 9–10) on the intention item. Values outside of parentheses represent frequency count, and values inside parentheses represent percentage of First Response.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

TABLE 3

FIML Fixed Parameter Estimates for Two-Level Growth Curve Models of Intention "to Pursue a Science-Related Research Career"

Fixed effects	**M1: Null**	**M2: + Time. Linear**	**M3: + Time. Quadratic**	**M4: + Match Status**	**M5: + Propensity Score**	**M7: + Research Experience**
For intercept						
Intercept (β_{00})	8.09 (.10)***	8.50 (.09)***	8.57 (.10)***	8.87 (.19)***	8.71 (0.19)***	8.73 (0.19)***
Match (β_{30})				−0.40 (.21)	−0.18 (.22)	−0.21 (0.22)
Propensity					2.37 (0.57)***	2.25 (0.57)***
score (β_{01})						
Time.Linear growth slope						
Intercept (β_{10})		−0.31 (.05)***	−0.57 (.12)***	−0.39 (.15)***	−0.39 (0.15)***	−0.74 (0.18)***
Time.Linear × Match (β_{40})				−0.24 (.11)***	−0.25 (.11)	−0.09 (.12)
Time.Quadratic growth			0.08 (.03)*	0.07 (.03)*	0.08 (.03)*	0.08 (.03)*
slope (β_{20})						
Research experience (β_{50})						0.43 (.13)***

Note. Values outside parentheses are fixed effects, values inside parentheses are standard error; M6: + Scientific Mentorship is not shown because it exhibited a nonsignificant effect on the outcome.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

TABLE 4

FIML Variance Components and Model Fit Estimates for Two-Level Growth Curve Models of Intention "to Pursue a Science-Related Research Career"

Variance components	M1: Null	M2: + Time. Linear	M3: + Time. Quadratic	M4: + Match status	M5: + Propensity score	M7: + Research experience
Within-Person (σ^2)	3.23	2.54	2.51	2.54	2.52	2.55
Intercept (τ_{00})	2.47 (.26)***	1.66 (.23)***	1.68 (.23)***	1.63 (.22)***	1.51 (.21)***	1.48 (.21)***
Linear (τ_{11})		0.32 (.06)***	0.32 (.06)***	0.28 (.05)***	0.29 (.05)***	0.26 (.05)***
Model Fit						
−2LL	5,263.06	5,141.83	5,136.31	5,121.14	5,104.19	5,093.03
AIC	5,269.06	5,151.83	5,148.31	5,137.14	5,122.19	5,113.03
Parameters	3	4	6	8	9	10

Note. −2LL = deviance statistic; AIC = Akaike's information criterion; parameters = number of parameters estimated in the model; values in parentheses are standard errors for the Wald test of statistical significance; M6: + Scientific Mentorship is not shown because it exhibited a nonsignificant effect on the outcome.

*** $p < .001$.