

# New Measures Assessing Predictors of Academic Persistence for Historically Underrepresented Racial/Ethnic Undergraduates in Science

Angela Byars-Winston,<sup>1\*</sup> Jenna Rogers,<sup>2</sup> Janet Branchaw,<sup>3</sup> Christine Pribbenow,<sup>4</sup> Ryan Hanke,<sup>5</sup> and Christine Pfund<sup>6</sup>

<sup>1</sup>Department of Medicine, Center for Women's Health Research, University of Wisconsin–Madison, Madison, WI 53715; <sup>2</sup>Center for Women's Health Research, University of Wisconsin–Madison, Madison, WI 53715; <sup>3</sup>WISCIENCE, Wisconsin Institute for Science Education and Community Engagement, University of Wisconsin–Madison, Madison, WI 53715; <sup>4</sup>Wisconsin Center for Education Research, University of Wisconsin–Madison, Madison, WI 53715; <sup>5</sup>Department of Educational Psychology, University of Wisconsin–Madison, Madison, WI 53715; <sup>6</sup>Department of Medicine, Institute for Clinical and Translational Research, Center for Women's Health Research, and Wisconsin Center for Education Research, University of Wisconsin–Madison, Madison, WI 53715

## ABSTRACT

An important step in broadening participation of historically underrepresented (HU) racial/ethnic groups in the sciences is the creation of measures validated with these groups that will allow for greater confidence in the results of investigations into factors that predict their persistence. This study introduces new measures of theoretically derived factors emanating from social cognitive and social identity theories associated with persistence for HU racial/ethnic groups in science disciplines. The purpose of this study was to investigate: 1) the internal reliability and factor analyses for measures of research-related self-efficacy beliefs, sources of self-efficacy, outcome expectations, and science identity; and 2) potential group differences in responses to the measures, examining the main and interaction effects of gender and race/ethnicity. Survey data came from a national sample of 688 undergraduate students in science majors who were primarily black/African American and Hispanic/Latino/a with a 2:1 ratio of females to males. Analyses yielded acceptable validity statistics and race × gender group differences were observed in mean responses to several measures. Implications for broadening participation of HU groups in the sciences are discussed regarding future tests of predictive models of student persistence and training programs to consider cultural diversity factors in their design.

## INTRODUCTION

There is a mismatch between the cultural demographics of the nation and the scientific workforce. This mismatch persists despite ongoing efforts to increase the participation of groups historically underrepresented (HU) in the sciences, including women and individuals from some racial and ethnic groups (Valantine and Collins, 2015). National efforts launched to diversify the scientific enterprise, ranging from the White House's 2009 *Educate to Innovate* initiative to the *America COMPETES Act* of 2007 and 2010, have all been targeted at increasing participation in career pathways in science, technology, engineering, and mathematics (STEM) for HU groups. Yet the progress of these groups earning science degrees, choosing science career paths, and entering the scientific workforce has been minimal (National Science Foundation [NSF], 2013). This minimal participation stands in sharp contrast to women, African Americans, Latino/as, and Asian Americans comprising both the majority of the U.S. population and, consequently, the largest

Kenneth Gibbs, *Monitoring Editor*

Submitted January 13, 2016; Revised May 25, 2016; Accepted May 25, 2016

CBE Life Sci Educ September 1, 2016 15:ar32

DOI:10.1187/cbe.16-01-0030

\*Address correspondence to: Angela Byars-Winston (ambyars@uwisc.edu).

© 2016 A. Byars-Winston et al. CBE—Life Sciences Education © 2016 The American Society for Cell Biology. This article is distributed by The American Society for Cell Biology under license from the author(s). It is available to the public under an Attribution–Noncommercial–Share Alike 3.0 Unported Creative Commons License (<http://creativecommons.org/licenses/by-nc-sa/3.0>).

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

share of new labor market entrants (Byars-Winston, 2014; Valantine and Collins, 2015).

In general, more than half of undergraduate first-year students from HU racial/ethnic groups, namely African Americans, Latino/as, and Native Americans, intend to major in a science or engineering degree (Hurtado *et al.*, 2007; President's Council of Advisors on Science and Technology, 2012). But together they earn only ~17% of bachelor's degrees in those fields (National Center for Education Statistics, 2014). Thus, underrepresentation in the sciences is not for lack of ability or interest (Seymour and Hewitt, 1997; Crisp *et al.*, 2009; Koenig, 2009). Underrepresentation, instead, appears to be a problem of translating the abilities and interests of students from racial/ethnic groups historically underrepresented in the sciences into persistence (Byars-Winston *et al.*, 2011).

### The Need for Measures Tested with Groups of Interest

A challenge in understanding factors that impact the persistence of individuals from HU racial/ethnic groups in the sciences is the dearth of theoretically based measures to assess their career and academic development experiences. Although research into the specific factors that account for key academic and career outcomes in the sciences is emerging, many of the published studies have not used measures validated for individuals in the sciences in general or for HU racial/ethnic groups in particular. Moreover, these studies have not reported the psychometric properties of measures used to operationalize their variables or based their measures in theory (Byars-Winston *et al.*, 2011; Trujillo and Tanner, 2014). To identify effective intervention strategies to increase science interests and persistence for individuals from HU racial/ethnic groups, we need measures that are theoretically informed, particularly by vocational and career theory, and psychometrically sound for these populations.

A central element of increasing college students' interest and persistence in science careers is their participation in undergraduate research experiences (President's Council of Advisors on Science and Technology, 2012). Extensive data support the positive impact of these experiences on students' self-reported gains in their research and writing skills, academic self-confidence, research productivity, and increased consideration of graduate degrees in science (see Laursen *et al.*, 2010). However, there is more to know about *how* such interventions exert their impact on student outcomes. For example, how do programmatic elements in the research training environment impact cognitive factors such as students' beliefs and identities?

### Purpose of the Study

In this study, we validated measures based on social cognitive and science identity theories with a sample of primarily black/African-American and Hispanic/Latino/a undergraduate students who were participating in undergraduate research. We introduced new measures of theoretically derived factors associated with student persistence, including self-efficacy and science identity, and examined potential variability in these factors as a function of gender, racial/ethnic group membership, and the intersection of gender and race/ethnicity. The findings from this study are important for professionals (e.g., instructors, mentors, program directors) working to increase the per-

sistence of HU racial/ethnic groups in science career pathways, because they provide tools to both monitor trainee research-related self-efficacy and science identity and to assess the impact of interventions. The findings will also contribute to future empirical research into the factors in students' scientific career and academic development that facilitate broadening participation for individuals from HU racial/ethnic groups.

## LITERATURE REVIEW

### Factors Influencing Persistence in Science

Career paths, including the science career path, are determined by a complex interaction of experiences and choices that begin in early childhood and are continually shaped by personal and environmental factors over the life course. Bakken *et al.* (2006) argued that current efforts to enhance persistence, particularly in clinical and biomedical sciences, largely focus on environmental elements (e.g., didactic training programs, financial support) and often ignore important person (e.g., race, gender) and cognitive factors (confidence, motivation, identity, etc.) that interact with research training environments. The interplay between a person's characteristics, his or her cognitions, and the research training environment shapes the person's career interests, choices, and adjustments (Graham *et al.*, 2013). Two cognitive factors that have been associated with persistence of undergraduates in science are self-efficacy and science identity (Trujillo and Tanner, 2014).

*The Self-Efficacy Factor.* One of the more well-researched cognitive factors in academic and career development literature is self-efficacy. Self-efficacy (confidence in one's ability to successfully perform a given task) is a central construct in social cognitive career theory (SCCT; Lent *et al.*, 1994) and is highly correlated with choice of and persistence in a science or engineering major (Lent *et al.*, 1986, 2005).

A model of SCCT that has been modified to incorporate science identity (discussed in the next section) is presented in Figure 1. The gray boxes in the figure indicate the variables measured in the present study. Briefly, SCCT posits that people are likely to form enduring *interest* in an activity when they view themselves as competent at performing it and when they expect the activity to produce valued outcomes (i.e., when they possess favorable *self-efficacy* and *outcome expectations* in relation to the activity). Along with self-efficacy and outcome expectations, personal interests foster particular educational and occupational *choice goals* (e.g., intentions to pursue a particular career path) that, in turn, make it more likely that people will take *actions* to achieve their goals (e.g., seek entry into particular academic or occupational paths). Individuals' subsequent *performance attainments* (e.g., successes, failures) provide valuable feedback that can strengthen or weaken self-efficacy and outcome expectations and, ultimately, help to revise or stabilize choices.

SCCT proposes that self-efficacy beliefs are formed via four types of information: performance accomplishments (e.g., personal mastery experiences or past successes), vicarious learning (e.g., observing the explicit behaviors of others, such as role models), social persuasion (e.g., verbal encouragement), and affective/emotional arousal experienced while completing a task (e.g., low anxiety, relaxed state). Importantly, the theory takes into account several pathways through which individual

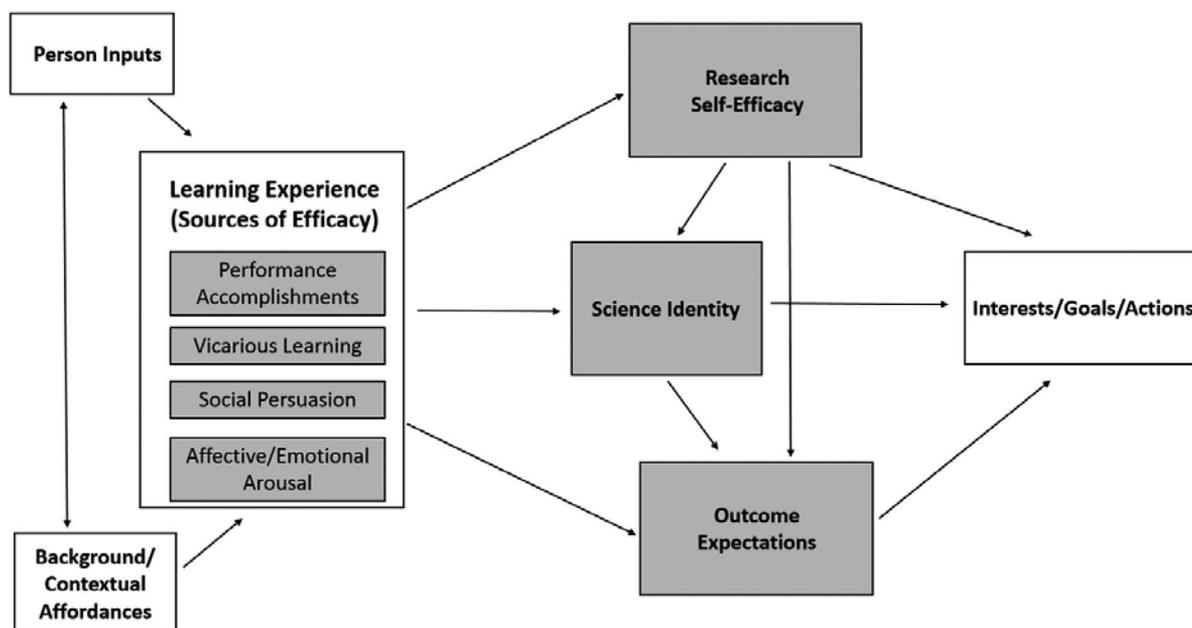


FIGURE 1. Modified model of SCCT incorporating science identity. Gray boxes indicate variables measured in the present study.

differences (e.g., gender, race, and personality) affect the academic and career development process.

Bakken *et al.* (2006) illustrated how SCCT, based on Bandura's (1997) social cognitive theory, provides a valuable theoretical base for understanding and promoting biomedical and clinical research careers, because it incorporates the reciprocal interactions between person, cognitive, and environmental elements in shaping career outcomes. Studies by Hurtado *et al.* (2009), Chemers *et al.* (2011), Estrada *et al.* (2011), Adedokun *et al.* (2012), and Byars-Winston *et al.* (2015) have focused on the contributions of self-efficacy factors to career outcomes for diverse students in the biological and biomedical sciences. Their research has shown support for scientific or research-related self-efficacy (confidence in one's ability to successfully perform scientific work or conduct research) as an important mediator between students' research experiences and their subsequent commitment to a research career.

**The Science Identity Factor.** All individuals have the fundamental need to belong (Allen and Eby, 2007). And that need to belong is often translated into an identification with a given group or context. Identity is defined as the point at which "an individual accepts influence from another person or a group in order to establish or maintain a satisfying self-defining relationship to the other" (Kelman, 2006, p. 3). Research demonstrates that students' identification (or sense of belonging, fit) with an academic setting is correlated with their academic success and persistence (see Gloria and Robinson Kurpius, 2001). Further, identification with identities that are relevant to a given context, like having a strong identity as an athlete while playing a sport, predicts performance in that context. When a student reports feeling like a scientist, and thereby assumes the identity of scientist, he or she is more likely to pursue a career in science (Estrada *et al.*, 2011). Research

shows that identifying as a scientist, or science identity, is related to persistence behaviors (e.g., applying for a PhD degree) (Seymour and Hewitt, 1997; Chemers *et al.*, 2011). Our modified SCCT model in Figure 1 places science identity as a potential intermediary factor, along with relevant self-efficacy and outcome expectancy beliefs, between students' learning experiences in research and their academic and career outcomes.

Science identity is informed by recognition of one's self and recognition by others as a potential scientist. Carlone and Johnson (2007) articulated a framework for science identity that includes three overlapping dimensions: *competence* (knowledge and understanding of science content), *performance* (skills and opportunities to act like a scientist), and *recognition* (characterized by acknowledging oneself and being recognized as a scientist by others). Choosing a science career means taking on the identity of a scientist. It is important to note that this science identity must also be negotiated along with other salient social identities (e.g., gender, race, class). As such, the ability to integrate multiple identities plays a role in persistence, too (Hurtado *et al.*, 2009; Laursen *et al.*, 2010; Chang *et al.*, 2011; Chemers *et al.*, 2011; Estrada *et al.*, 2011; Thew *et al.*, 2012).

Studies examining science identity and how it operates to contribute to student academic and career outcomes, particularly for HU racial/ethnic groups, are relatively recent, and there is still more to be learned. For instance, how does a research experience or a research mentor impact a student's science identity and does science identity moderate those influences on a student's subsequent science career commitment? Emerging findings from studies by Estrada *et al.* (2011) and Chemers *et al.* (2011) suggest that science identity plays a complementary role along with science or research-related self-efficacy beliefs in driving students' science academic intentions (e.g., intentions to complete a baccalaureate or graduate

degree). Placement of science identity within our modified SCCT model in Figure 1 is based on the findings of Estrada *et al.* (2011) and Chemers *et al.* (2011).

### Contribution of This Study to Broadening Science Participation

There is a published body of work with theoretically informed measures that assess the experiences, attitudes, and beliefs of diverse undergraduate students in science, particularly the work of Chemers *et al.* (2011) and Estrada *et al.* (2011). We build and expand upon that important work in several ways. First, whereas the existing science self-efficacy measures assess one's confidence to perform various research and science-related tasks, we expand our research self-efficacy measure to capture both confidence in performing research tasks and persisting in a research science career pathway (i.e., complete undergraduate degree, pursue and complete graduate degree), following the SCCT work of Byars-Winston *et al.* (2010) and Lent *et al.* (2005). Second, although there are validity data for several extant measures of science self-efficacy and science identity, we could not find data on the measurements' performance (i.e., invariance) with different cultural groups. In this study, we examine validity data regarding our measures' functioning across diverse groups. Finally, no measures of sources of self-efficacy in a research domain exist. Toward the advancement of full tests of the SCCT model (Lent *et al.*, 1994), such as the one depicted in Figure 1, we developed and report on a new measure of research-related sources of self-efficacy in the present study.

## METHOD

### Participants

Since 2001, the American Society for Microbiology (ASM) has sponsored the Annual Biomedical Research Conference for Minority Students (ABRCMS) with continuing funds from the National Institutes of Health (NIH). This conference was designed with the dual purpose of encouraging undergraduate students to pursue advanced training and careers in the biomedical sciences and providing faculty mentors, advisors, and program leaders with resources for facilitating student success. In particular, the specific aims of ABRCMS ([www.abrcms.org](http://www.abrcms.org)) are for undergraduate students to

- exchange their research findings and demonstrate scientific expertise;
- prepare for the evolving, global, and interdisciplinary nature of biomedical and behavioral sciences research; and
- transition successfully *from* undergraduate, postbaccalaureate, and master's programs *to* graduate or professional education.

Undergraduate and postbaccalaureate student attendance has doubled since the first year of the conference (2001 = 976; 2013 = 1775), and students consistently represent ~52% of the total conference attendees.

### Data Collection

Undergraduate and postbaccalaureate students, as well as graduate student and faculty/staff attendees of ABRCMS, are asked to complete a postconference survey on the final day of the conference. The survey is primarily used for formative pur-

poses; hence, the questions are geared toward evaluating satisfaction with the conference and conference logistics and content. For better assessment of students' experiences and outcomes, additional items that were approved by the University of Wisconsin–Madison Institutional Review Board (IRB) for Human Subjects Research were added to the undergraduate and postbaccalaureate survey in 2012 and 2013. The surveys were sent electronically using Qualtrics to everyone who had registered for the conference in those 2 yr. They were given approximately 2 wk to complete the survey, with a reminder sent once during the 2-wk period. In 2012, 508 students completed the survey for a response rate of 28%; 666 completed it in 2013 for a response rate of 38%. From these 1174 responses, any participant who did not opt into research ( $n = 400$ ) or who provided incomplete data ( $n = 86$ ) was removed before the analysis, leaving a final sample of 688.

### Instrument/Survey Development

Item content was generated via two processes. First, items were adapted from existing SCCT measures (Lent *et al.*, 2005; Byars-Winston *et al.*, 2010), social cognitive theory (Usher and Pajares, 2006), a mentee survey used in a summer research opportunity program (Pfund *et al.*, 2006; Byars-Winston *et al.*, 2015), and the Undergraduate Research Student Self-Assessment (URSSA; Weston and Laursen, 2015) to create the research self-efficacy and the sources of self-efficacy items. Second, additional items were generated to capture further learning experiences based on expert opinion from a research team member (J.B.) who is a faculty member and a program director leading a national undergraduate research program. The expert opinion was particularly informative to the development of sources of self-efficacy items, because, to the team's knowledge, there have been few published sources of self-efficacy scales in the STEM domain since the first one by Lent *et al.* (1991), and only one has been used with participants from an HU racial/ethnic group, and this was limited to assessing sources of self-efficacy for mathematics (Gainor and Lent, 1998). These adapted and newly generated items were used to create four measures: research self-efficacy, sources of self-efficacy, outcome expectations, and science identity, which are described in greater detail in the following sections. A complete list of items for each of the four measures appears in Table 1.

**Research Self-Efficacy.** The research self-efficacy scale includes item content that covers academic and career-related benchmarks of efficacy. These items assess individuals' perceptions about their ability to both perform and persist in science domains, consistent with self-efficacy item content in earlier measures by Byars-Winston *et al.* (2010) and Lent *et al.* (2005), whose science or engineering self-efficacy measures evidenced reliability coefficients of  $\alpha = 0.91$ – $0.92$ . Participants were instructed to rate themselves on a five-point scale (ranging from 1 = no confidence to 5 = complete confidence), indicating their degree of confidence in their ability to perform or complete 11 tasks related to doing research and attainment of science undergraduate and graduate degrees.

**Sources of Self-Efficacy.** The sources of self-efficacy scale includes 10 items participants had encountered in their most

TABLE 1. Scale items

<b>Research self-efficacy</b>
How much confidence do you have in your ability to:
Make important contributions to a research team?
Explain your research topic to other scientists?
Do research?
Write a basic research proposal?
Present a research talk or poster?
Excel in your science major over the next two semesters?
Pursue a research science career?
Complete a science degree?
Persist with science courses even though you may be a minority in them?
Pursue a graduate degree in science?
Complete a graduate degree in science?
<b>Sources of self-efficacy</b>
Performance accomplishments
Based on feedback from mentor, in last research experience how well did you:
Independently conduct experiments or a research project?
Analyze research data?
Write a scientific report?
Prepare a scientific poster or presentation?
Vicarious learning
My primary research mentor showed me how to conduct a research procedure.
I look up to my research mentor as a career role model.
Social persuasion
My research mentor encouraged me to pursue a research science career.
My research mentor told me I have the ability to be a scientist.
Affective/emotional arousal
I felt nervous when conducting research.
I felt anxious about my ability to do research.
<b>Outcome expectations</b>
A research science career would allow me to:
Do work that makes a difference in people's lives or society.
Do work that I find satisfying.
Go into a field with high employment demand.
Get respect from other people.
Earn an attractive salary.
<b>Science identity</b>
During my most recent research experience, I:
Felt like a scientist.
Interacted with scientists from outside of my school.
Felt part of a scientific community.

recent research experience, with content relating to performance accomplishments (four items), vicarious learning (two items), social persuasion (two items), and affective/emotional arousal (two items). For the performance accomplishments items, participants were instructed to rate how well they had carried out each of the tasks (ranging from 1 = not well at all to 5 = extremely well), and for the remaining domains of items, they rated the degree to which they agreed they had encountered each of the activities (ranging from 1 = strongly disagree to 5 = strongly agree). We acknowledge that context has a large

impact on college student experiences and can be variable across both time and across student groups, with greater variability and inconsistency of supports and opportunities for HU students (e.g., Carter, 2006). Because of that variability and inconsistency, we chose to focus on participants' most recent research experience as opposed to the sum of all their research experiences up to the point of the study in order to assess sources of self-efficacy in one particular context.

**Outcome Expectations.** We used a modified version of the outcome expectations scale from Lent *et al.* (2005), for which a reliability coefficient of  $\alpha = 0.89$  was observed with a sample of racially diverse engineering undergraduate students. We chose five items that captured the three types of outcome expectations defined in SCCT (Lent *et al.*, 1994): physical (e.g., monetary), self-evaluative (e.g., personal satisfaction), and social (e.g., respect from others). Participants were instructed to rate the degree to which they agreed with the various benefits that a research science career would personally provide them, ranging on a scale from 1 = strongly disagree to 5 = strongly agree.

**Science Identity.** The seven items in the science identity scale were drawn from the Attitudes and Behaviors as a Researcher component of the URSSA tool and included practicing authentic scientific inquiry and feeling responsible for a research project (Weston and Laursen, 2015). Weston and Laursen (2015) reported reliability coefficients for this URSSA component of  $\alpha = 0.83$ – $0.84$  with a national sample of undergraduate students involved in research. We selected the three items in the component that directly assessed individuals' interacting with the scientific community and feeling like a scientist. Participants rated their strength of agreement with the items on a five-point scale, ranging from 1 = strongly disagree to 5 = strongly agree. Following our rationale and decision for the context of the sources of self-efficacy scale, participants were instructed to focus on their most recent mentored research experience when responding to the science identity items.

#### Data Analysis for Validation of Scales

The first set of analyses sought to validate the scales with a sample of undergraduate students in the sciences consisting of mainly students from HU racial/ethnic groups. There are four stages of scale development (Netemeyer *et al.*, 2003). The first is the construct definition through theoretical and literature review, the second is the generation of the measurement items, the third is the refinement of scales through testing, and the last is the finalizing of the scales. Stages 1 and 2 are described above, while stages 3 and 4 are covered in the psychometric analyses presented in this paper. Specifically, these analyses explored the internal consistency, and construct validity of each of the four scales (research self-efficacy, sources of self-efficacy, outcome expectations, and science identity).

Scale construct validity and internal consistency were examined through a series of confirmatory factor analyses (CFAs) and use of coefficient alpha for the entire sample of undergraduates. Next, we examined the construct validity as well as the internal consistency of each scale on the individual groups (black/African-American females, black/African-American males, Hispanic/Latino males, and Hispanic/Latina females) to

**TABLE 2. Results of CFAs for full sample**

	Mean <sup>a</sup>	SD	Alpha	RMSEA	SRMR	CFI
Research self-efficacy	4.25	0.660	0.910	0.173	0.099	0.831
Sources of self-efficacy	4.14	0.841	0.673	0.061	0.044	0.960
Outcome expectations	4.32	0.570	0.788	0.079	0.028	0.986
Science identity	4.28	0.703	0.730	—	—	—

<sup>a</sup>Range: 1–5.

ensure that the factor structure and internal consistency were acceptable across these diverse groups representing race/ethnicity and gender intersections.

For the initial psychometric analyses, the entire sample of survey respondents was included ( $N = 688$ ). Of the 688 participants included in this study, there was almost a 2:1 ratio of females ( $n = 458$ ; 67%) to males ( $n = 230$ ; 33%). Participants self-reported their racial/ethnic groups as follows: black/African American ( $n = 288$ , 42%), Asian American ( $n = 53$ , 8%), Hispanic/Latino(a) ( $n = 275$ , 40%), Native American ( $n = 7$ , 1%), and Pacific Islander/Alaskan Native ( $n = 5$ , 0.7%). A smaller number of participants identified as Caucasian ( $n = 28$ , 4%) and other ( $n = 32$ , 4%). A subsample, described in the following paragraph, was used for the individual group analyses.

#### Data Analyses for Testing Group Differences

Potential group differences by gender and race/ethnicity were examined as was the intersection of the two statuses. We performed  $t$  tests and analysis of variance (ANOVA) to test for differences between the groups. Results and details of these analyses are described below. Large sample sizes allowed a second set of analyses examining group differences on mean scale scores to be performed. A subsample of black/African-American and Hispanic/Latino(a) respondents ( $N = 563$ ) was evaluated to explore potential racial/ethnic and gender differences in the mean response scores on the scales. The subsample included 197 black/African-American females, 91 black/African-American males, 175 Hispanic/Latina females, and 100 Hispanic/Latino males.

## RESULTS

### Psychometric Analyses Validate the Scales on a Sample of Undergraduates from Predominantly HU Racial/Ethnic Groups

In this section, we present the results of the analyses exploring the construct validity and the internal consistency of each of the four scales (research self-efficacy, sources of self-efficacy, outcome expectations, and science identity). Results are presented first for the full sample and second for the individual demographic groups.

#### CFAs Verify Scale Validity

A series of CFAs were run to verify scale validity. These analyses were completed using a maximum-likelihood estimation procedure in the EQS statistical program. We used a combination of measurement properties to estimate and assess model fit using comparative fit statistic (CFI), the standardized root-mean-square residual (SRMR), and the root-mean-square error of approximation (RMSEA; McDonald, 1989; Hu and Bentler, 1999; Jackson et al., 2009). Models demonstrating an RMSEA

greater than 0.10 indicate a poor fit, whereas models with an SRMR of less than 0.08 and a CFI value greater than 0.90 indicate a strong fit (Hu and Bentler, 1999). It was not possible to run a CFA with the science identity scale due to the low number of items (three), which resulted in a model with 0 degrees of freedom that did not permit fit statistics to be calculated. An exploratory factor analysis was conducted as an alternate approach to look at the factor structure of the science identity scale.

#### Full Sample

Results of the CFA for the full sample along with descriptive statistics appear in Table 2. The research self-efficacy model demonstrated an adequate fit, with RMSEA = 0.173, SRMR = 0.099, and CFI = 0.831. There is evidence of a strong fit for the sources of self-efficacy scale, with RMSEA = 0.061, SRMR = 0.044, and CFI = 0.960. The model for the outcome expectations scale demonstrated a strong fit, RMSEA = 0.079, SRMR = 0.028, and CFI = 0.986 (Jöreskog and Sörbom, 1982; Browne and Cudeck, 1993; Hu and Bentler, 1999). We also ran an exploratory factor analysis on the science identity scale for each of the four groups at the intersection of race/ethnicity and gender in our study. Results indicate that the scale items all load on one factor for each of the groups, thus demonstrating that the science identity factor structure is similar across the four groups.

#### Individual Demographic Groups

We next examined the factor structure of the research self-efficacy, sources of self-efficacy, and outcome expectations scales among four different demographic groups: black/African-American males, black/African-American females, Hispanic/Latino males, and Hispanic/Latina females. Table 3 provides the results for the CFAs conducted on each of the demographic groups.

**Research Self-Efficacy.** The results show that the research self-efficacy scale has a weak or adequate fit for each of the individual groups examined in this study. For black/African-American males, the CFA models for the research self-efficacy scale demonstrated poor fit (RMSEA = 0.222, SRMR = 0.082, and CFI = 0.772). This was also the case for the CFA model including black/African-American females (RMSEA = 0.203, SRMR = 0.098, and CFI = 0.799). The CFA model for the research self-efficacy scale that included Hispanic/Latino males demonstrated poor fit (RMSEA = 0.146, SRMR = 0.103, and CFI = 0.853), as did the model including Hispanic/Latina females (RMSEA = 0.142, SRMR = 0.117, and CFI = 0.868).

**Sources of Self-Efficacy.** CFA models examining the sources of self-efficacy scale demonstrated overall good fit for the

TABLE 3. Results of confirmatory factor analyses by demographic groups

	Black/African-American males				Black/African-American females				Hispanic/Latino males				Hispanic/Latina females			
	Alpha	RMSEA	SRMR	CFI	Alpha	RMSEA	SRMR	CFI	Alpha	RMSEA	SRMR	CFI	Alpha	RMSEA	SRMR	CFI
Research self-efficacy	0.927	0.222	0.082	0.772	0.921	0.203	0.098	0.799	0.869	0.146	0.103	0.853	0.627	0.142	0.117	0.868
Sources of self-efficacy	0.730	0.106	0.100	0.906	0.711	0.050	0.048	0.974	0.719	0.082	0.060	0.937	0.886	0.086	0.067	0.924
Outcome expectations	0.790	0.041	0.029	0.996	0.816	0.134	0.054	0.968	0.716	0.052	0.037	0.992	0.771	0.115	0.044	0.966
Science identity	0.783	—	—	—	0.701	—	—	—	0.687	—	—	—	0.719	—	—	—

individual groups, with the exception of the model that included black/African-American males. For black/African-American males, the CFA model for the sources of self-efficacy scale demonstrated adequate fit (RMSEA = 0.106, SRMR = 0.100, and CFI = 0.906). For black/African-American females, the CFA model for the sources of self-efficacy scale was a strong fit (RMSEA = 0.050, SRMR = 0.048, and CFI = 0.974). The CFA model for the sources of self-efficacy scale that included Hispanic/Latino males had an RMSEA = 0.082, SRMR = 0.060, and CFI = 0.937, all of which indicate a strong fit. This was also the case for the CFA model including Hispanic/Latina females (RMSEA = 0.086, SRMR = 0.067, and CFI = 0.924).

**Outcome Expectations.** Model fit statistics for the outcome expectations scale indicate that the scale has a good fit across each of the four groups we examined, though the model was a better overall fit for black/African-American males and Hispanic/Latino males than for the complementary female samples used in this study. For black/African-American males the CFA models for the outcome expectations scale demonstrated good fit (RMSEA = 0.041, SRMR = 0.029, and CFI = 0.996). The CFA model for the outcome expectations scale also demonstrated good fit for black/African-American females in our sample (RMSEA = 0.134, SRMR = 0.054, and CFI = 0.968). For Hispanic/Latino males, the CFA for the outcome expectations scale demonstrated a strong fit (RMSEA = 0.052, SRMR = 0.037, and CFI = 0.992). The model examining the factor structure of the outcome expectations scale with Hispanic/Latina females had a good fit (RMSEA = 0.115, SRMR = 0.044, and CFI = 0.966).

To summarize, our analyses of each scale via CFA revealed evidence of construct validity among this group of predominantly HU students. Evidence of construct validity was also found for each of the scales for our four demographic groups, with the exception of the research self-efficacy scale.

#### Coefficient Alpha Tests Internal Consistency

For the last step of the psychometric analyses, a series of internal consistency estimates were run using the coefficient alpha ( $\alpha$ ), which is widely used to establish levels of internal consistency (Netemeyer *et al.*, 2003). This statistic is representative of the consistency (i.e., relationship) between items in a given scale. Values range from 0 to 1, with anything 0.70 and greater being generally considered as good and showing internal consistency, and values between 0.60 and 0.70 being acceptable for parsimonious measurement instruments and scale-development purposes (Netemeyer *et al.*, 2003).

**Full Sample.** Internal consistency statistics for the full sample for each of the four scales are presented in Table 2. The alpha for the research self-efficacy scale was 0.910, demonstrating very strong reliability. The alpha for the sources of self-efficacy scale was 0.673, which is bordering on acceptable. However, this perhaps is not surprising, given that four different dimensions (types of sources of self-efficacy) are being measured with this scale. The alphas for the outcome expectations scale and the science identity scale were 0.788 and 0.730, respectively, both indicating good internal consistency.

**Individual Demographic Groups.** Internal consistency statistics for the research self-efficacy, sources of self-efficacy,

outcome expectations, and science identity scales are reported in Table 3. The alpha coefficient for research self-efficacy is good across all groups, with the exception of Hispanic/Latina females, for which it is fairly low ( $\alpha = 0.627$ ). The alpha coefficients for the sources of self-efficacy scale are good across all groups, though the scale demonstrates the highest level of internal consistency among Hispanic/Latina females. Findings show that the alpha coefficient for outcome expectations is good across all four groups, with the highest level of internal consistency found among black/African-American male participants. Internal consistency statistics for the science identity scale were good across all four groups ( $\alpha_{\text{black/African-American males}} = 0.783$ ,  $\alpha_{\text{black/African-American females}} = 0.701$ ,  $\alpha_{\text{Hispanic/Latino males}} = 0.687$ ,  $\alpha_{\text{Hispanic/Latina females}} = 0.719$ ).

**Tests of Group Differences**

The CFA results demonstrated that the factor structure of the sources of self-efficacy and outcome expectations scales hold up within the whole group and within the different populations. The research self-efficacy scale, however, only demonstrated an adequate fit across the whole group and a poor to adequate fit across the individual demographic groups. We also found that the science identity scale is valid for these groups through the use of an exploratory factor analysis. A consistent factor structure means that we have the same variables across different groups. Despite this consistency, there can still be differences between the groups. To determine these differences across different racial/ethnic groups and genders, we performed a series of *t* tests (gender differences) and ANOVA tests (racial/ethnic differences).

The sources of self-efficacy scale was divided into four subscales to evaluate how each of the different groups responded to the sources. The results of these analyses are reported in Table 4. There were no significant differences in mean response levels between males and females for any of the scales. However, findings show that there are significant differences in the mean response levels between black/African-American and Hispanic/Latino(a) respondents on the affective/emotional arousal scale and the science identity scale. Further inspection of the intersection of race/ethnicity and gender revealed statistically significant differences, with a higher mean response on the affective/emotional arousal scale for black/African-American males and a higher mean response on the science identity scale for Hispanic/Latina females compared with the remaining groups.

**DISCUSSION**

Our goal in this study was to use SCCT and science identity theory to validate measures assessing persistence-related factors with students from HU racial/ethnic groups in science majors engaged in research experiences. Students from HU groups are the focus of much discussion regarding broadening, and by “broadening” we mean racially and ethnically diversifying, the talent pool of emerging scientists. We contend that an important step in the science of broadening participation of HU groups in the sciences is the creation and use of valid measures that will allow for greater confidence in the results of investigations into persistence factors for these groups. In considering the contributions of this study, we organize our discussion

**TABLE 4. Testing group differences on scales between black/African-American and Hispanic/Latino(a) participants (n = 563)**

	Gender		Race		Gender x Race				f
	Male	Female	Black/African American	Hispanic/Latino(a)	Black/African-American male	Black/African-American female	Hispanic/Latino male	Hispanic/Latino female	
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	
Research self-efficacy	4.26 (0.755)	4.25 (0.637)	4.19 (0.724)	4.33 (0.621)	4.22 (0.735)	4.17 (0.720)	4.30 (0.775)	4.35 (0.514)	2.39
Sources of self-efficacy									
Performance accomplishments	4.09 (0.908)	4.16 (0.835)	4.15 (0.844)	4.16 (0.878)	4.07 (0.849)	4.13 (0.843)	4.11 (0.962)	4.18 (0.828)	0.381
Vicarious learning	4.30 (0.703)	4.22 (0.766)	4.17 (0.774)	4.33 (0.707)	4.22 (0.659)	4.15 (0.822)	4.38 (0.736)	4.30 (0.691)	2.48
Social persuasion	4.28 (0.775)	4.19 (0.819)	4.14 (0.855)	4.31 (0.742)	4.22 (0.779)	4.11 (0.887)	4.35 (0.771)	4.29 (0.726)	2.51
Affective/emotional arousal	3.23 (1.06)	3.09 (1.20)	3.24 (1.14)	3.02 (1.16)	3.33 (1.02)	3.20 (1.20)	3.13 (1.10)	2.95 (1.19)	2.59*
Outcome expectations	4.31 (0.544)	4.33 (0.567)	4.33 (0.582)	4.31 (5.33)	4.30 (0.607)	4.35 (0.571)	4.32 (0.482)	4.31 (0.561)	0.243
Science identity	4.25 (0.758)	4.30 (0.693)	4.22 (0.754)	4.35 (0.668)	4.25 (0.809)	4.21 (0.729)	4.25 (0.712)	4.41 (0.636)	2.76*

\*p < 0.05.

around three primary points: measurement matters, group differences, and implications of the results for practitioners in designing and evaluating programmatic interventions.

### Performance of Measures

We have provided a new set of measures to assess HU racial/ethnic students' beliefs about their personal research-related self-efficacy, outcome expectations, and science identities. Analyses of these measures yielded acceptable validity and reliability statistics. Although other studies have used measures of self-efficacy and, to a lesser degree, outcome expectations, we developed the first scale to assess research-related sources of efficacy. This scale now allows for future study of learning experiences that give rise to research self-efficacy beliefs and the learning experiences that give rise to outcome expectations following Fouad and Guillen's (2006) call for more studies of predictive factors for outcome expectations.

The science identity scale exhibited modest but acceptable reliability when tested on our primarily HU racial/ethnic group. While these items were used as a scale cluster in previous studies, they have not been validated on a group of HU undergraduate students in science disciplines. There is room for improvement on this scale. New items could be tested and added to more accurately capture "science identity" for HU racial/ethnic students and subsequently improve reliability. Further validity and reliability studies of the scale with other samples of HU racial/ethnic students and students who are not from HU backgrounds are needed. Such investigations would help to clarify whether the science identity scale items are more or less reliable for all groups of students or whether there is something particular about how the items operate for HU racial/ethnic groups of students.

Interestingly, we found that two sources of efficacy (social persuasion and vicarious learning) loaded onto one factor. Lent *et al.* (1994) theorized social persuasion (e.g., verbal encouragement) and vicarious learning (e.g., observing the explicit behaviors of others, such as role models) to be two unique sources contributing to an individual's self-efficacy beliefs. However, as our findings demonstrated, the questions intending to measure these two separate sources captured the same underlying factor. Thus, for the individuals in our study, these questions were not distinguishable in their minds, but rather these two sources seem to be conflated into one. That is, observing the behaviors of role models and receiving verbal encouragement operate together or act as a combined source of learning for these individuals. This makes sense, given that mentored research experiences are saturated with observational learning (Laursen *et al.*, 2010), and our data suggest that such learning is potentiated by socially persuasive and encouraging comments. It could also mean that, in the experiences of these individuals, verbal encouragement is only salient when coming from role models they have been observing. By extension, our findings suggest that, for this sample, learning research by observing others model research tasks co-occurs with the verbal encouragement students receive from those they observe. Research mentors would do well to be cognizant that they offer positive encouragement while supporting students' acquisition and mastery of research skills and tasks.

Overall, the results of our validity and reliability analyses indicate that these scales are appropriate for measuring

research-related sources of self-efficacy, outcome expectations, and science identity with each of the four demographic groups we examined. Our scales can facilitate future tests of the expanded SCCT model we proposed in Figure 1, which incorporates science identity, examining how the model variables work in relationship with desired student outcomes for HU racial/ethnic students in the sciences. The results from future examinations of the variables can inform mentor interventions, such as identifying what factors build mentees' research self-efficacy and then training mentors to address those factors in their research-mentoring relationships.

Relative to the research self-efficacy scale, the internal reliability coefficients on the total sample and the four demographic groups indicate that items in this scale are consistently measuring the same construct. This is useful information, as we purported to assess both students' perceptions of their ability to perform research tasks and persist in a research science career pathway. However, our examination of measurement invariance, or how the measure performed across groups, revealed that the scale better captures the content of research self-efficacy beliefs of the black/African-American males and females and the Hispanic/Latino males in our study, less so for Hispanic/Latina females. The observed difference for Hispanic/Latina females on the research self-efficacy scale both indicates that more item development and testing are warranted to increase the scale's construct validity and illustrates the importance of future studies examining group differences in their analyses of measures. Our findings also demonstrate that there is variation in the way that each different demographic group experiences the variables assessed in the measures. By extension, our findings also suggest that the influence of the mentored research experience as measured by these four scales may not be experienced in the same way by individuals from these four demographic groups, a point discussed further in the next section.

### Group Differences

While we validated the scales in this study using a group that was composed of all racial/ethnic groups (though predominantly HU groups), we took advantage of the size of the data set, which allowed for the examination of both between-group differences, testing the main effects for race/ethnicity and gender, and within-group differences, examining intersection effects of race/ethnicity with gender. In evaluating mean responses to items for black/African-American males and females and Hispanic/Latino/a males and females, we found that black/African-American males reported significantly higher negative affective/emotional arousal for doing research than the other three groups, and Hispanic/Latina females reported significantly higher science identity.

There may be multiple reasons why these two groups emerged with significant mean differences from their counterparts. It may be that the same research training environment is either experienced differently by black/African-American males and Hispanic/Latina females or that the research training environment itself provides different experiences to various groups. Black/African-American men are the only group among HU racial/ethnic groups in STEM fields not making progress in degree attainment (Bidwell, 2015). Their percentage of science and engineering bachelor's degrees has remained essentially

the same in the past decade, from 6.1% in 2002 to 6.2% in 2012 (NSF, 2013). Lack of critical mass, lack of “relatability” (Bidwell, 2015), and being a “minority among minorities” can lead to concerns over potentially fulfilling stereotypes about one’s group (i.e., stereotype threat), all of which can increase one’s anxiety about one’s performance. Hispanic/Latina females, conversely, are not only part of the fastest-growing ethnic groups in the United States but receive more baccalaureate degrees in science and engineering than Hispanic/Latino males (56 vs. 44%; NSF, 2013). Some studies have found that Hispanic/Latina females who are raised in a predominantly patriarchal household are more likely to consider science careers, perhaps owing to fathers’ encouragement of their daughters to consider “male” career options (Bowman, 1993; Marín, 1993). Whether or not Hispanic/Latina females begin their research experiences with higher science identity or elements of the research training environment (e.g., mentors, peers) support and nurture their science identity to a greater degree than it does for other students is unclear.

Overall, our results highlight the gendered nature of racial experiences in science (Lundy-Wagner, 2013). The emergence of group variation in mean responses to the study’s scales suggests that our measures are sensitive enough to detect subtle differences between racial/ethnic and gender groups, specifically for science identity and the sources of self-efficacy. The group differences should be interpreted with caution. Although the results were statistically significant, the present analyses do not provide any data regarding the consequence or impact of these group differences on other variables or outcomes for students. Further investigations can determine whether these differences are replicated with other samples of HU racial/ethnic students in the sciences. Additionally, different analytical methods, including multivariate analyses, can examine whether group differences are also present in how the variables correlate with each other. For example, how do the research-related sources of self-efficacy differentially explain the variance in research-related self-efficacy and outcome expectancy beliefs? Do these sources of self-efficacy also significantly contribute to science identity?

### Implications for Practitioners

The scales we developed and report on herein are important tools for practitioners, such as research program directors and research-based laboratory instructors, who aim to broaden participation in research and the scientific workforce. Practitioners need assessment and evaluation tools that generate valid data about the impact and outcomes of their programmatic and classroom interventions in order to determine whether they are effective for the targeted populations of students. In particular, the nuanced differences in responses between race/ethnicity and gender groups revealed by the analyses presented in this paper reflect important differences in experiences that must be considered when designing and revising interventions.

Beyond the scarcity of evaluation tools available for use by practitioners that have been validated with students from HU racial/ethnic groups, many practitioners, in particular those not trained in the social sciences, are unable to easily select appropriate assessment and evaluation tools. A first step to address this, taken here, is to publish scale development and validation research in journals that reach practitioners and that provide

guidance regarding how these tools may be used in the field. Interdisciplinary collaborations between life and social scientists are needed to advance our common goal of broadening participation in science. The research presented here reflects that type of collaboration.

### Limitations

The item content captures self-reported ratings of individuals’ beliefs about their abilities, experiences, and identities. They were not designed to be used as formal assessments of students’ abilities and experiences. Whereas the scales for research self-efficacy, outcome expectations, and science identity are worded such that they can be administered as pre or post measures, the content and wording for the sources of efficacy relate to experiences encountered during a specific research training event and, thus, are suited for posttraining measures.

### CONCLUSION

There is an increased need for scientific approaches to identifying, implementing, and evaluating initiatives that can advance the academic and career attainment of HU groups in the sciences (Valantine and Collins, 2015). Scientific inquiry into a phenomenon requires quantitative as well as qualitative tools that are valid for the populations of interest experiencing the phenomenon so that there is confidence in the findings. The measures in this study provide tools for investigating the research learning experiences of HU racial/ethnic students and how those experiences contribute to their persistence in science, and information for practitioners to consider about group differences when designing programs.

### ACKNOWLEDGMENTS

We gratefully acknowledge the comments and support from Anne-Barrie Hunter, Robert W. Lent, and Nancy Thayer-Hart and help from ABRCMS staff, Clifford W. Houston, and Amy L. Chang. This study was supported with funding from the NIH under grant R01 GM094573 and funding from the Department of Medicine at the University of Wisconsin–Madison. This is independent research, and the views expressed here do not indicate endorsement by the sponsors.

### REFERENCES

- Adedokun OA, Zhang D, Parker LC, Bessenbacher A, Childress A, Burgess WD, Adedokun BOA (2012). Understanding how undergraduate research experiences influence student aspirations for research careers and graduate education. *J Coll Sci Teach* 42, 82–90.
- Allen TD, Eby LT (2007). Common bonds: an integrative perspective on mentoring. In: *The Blackwell Handbook of Mentoring: A Multiple Perspectives Approach*, ed. T Allen and L Eby, Oxford, UK: Blackwell.
- Bakken LL, Byars-Winston A, Wang M-F (2006). Viewing clinical research career development through the lens of social cognitive career theory. *Adv Health Sci Educ Theory Pract* 11(1), 91–110.
- Bidwell A (2015 May 7). African-American men: the other STEM minority. *US News and World Report*. [www.usnews.com/news/stem-solutions/articles/2015/05/07/african-american-men-the-other-stem-minority](http://www.usnews.com/news/stem-solutions/articles/2015/05/07/african-american-men-the-other-stem-minority).
- Bowman SL (1993). Career intervention strategies for ethnic minorities. *Career Dev Q* 42, 14–25.
- Browne M, Cudeck R (1993). Alternative ways of assessing model fit. *Sociol Method Res* 21, 230–258.
- Byars-Winston A (2014). Toward a framework for multicultural STEM-focused career interventions. *Career Dev Q* 62, 340–357.

- Byars-Winston AM, Branchaw J, Pfund C, Leverett P, Newton J (2015). Culturally diverse undergraduate researchers' academic outcomes and perceptions of their research mentoring relationships. *Int J Sci Educ* 37, 2533–2554.
- Byars-Winston A, Estrada Y, Howard C, Davis D, Zalapa J (2010). Influence of social cognitive and ethnic variables on academic goals of underrepresented students in science and engineering: a multiple-groups analysis. *J Couns Psychol* 57, 205–218.
- Byars-Winston A, Gutierrez B, Topp S, Carnes M (2011). Integrating theory and practice to increase scientific workforce diversity: a framework for career development in graduate research training. *CBE Life Sci Educ* 10, 357–367.
- Carlone H, Johnson A (2007). Understanding the science experiences of successful women of color: science identity as an analytic lens. *J Sci Teach* 44, 1187–1218.
- Carter DF (2006). Key issues in the persistence of underrepresented minority students. *New Dir Inst Res* 130, 33–46.
- Chang MJ, Eagan MK, Lin MH, Hurtado S (2011). Considering the impact of racial stigmas and science identity: persistence among biomedical and behavioral science aspirants. *J High Educ* 82, 564–596.
- Chemers MM, Zurbriggen EL, Syed M, Goza BK, Bearman S (2011). The role of efficacy and identity in science career commitment among underrepresented minority students. *J Soc Issues* 67, 469–491.
- Crisp RJ, Bache LM, Maitner AT (2009). Dynamics of social comparison in counter-stereotypic domains: stereotype boost, not stereotype threat, for women engineering majors. *Soc Influence* 4, 171–184.
- Estrada M, Woodcock A, Hernandez PR, Schultz PW (2011). Toward a model of social influence that explains minority student integration into the scientific community. *J Educ Psychol* 103, 206–222.
- Fouad NA, Guillen A (2006). Outcome expectations: looking to the past and potential future. *J Career Assessment* 14, 130–142.
- Gainor KA, Lent RW (1998). Social cognitive expectations and racial identity attitudes in predicting the math choice intentions of black college students. *J Couns Psychol* 45, 403–413.
- Gloria AM, Robinson Kurpius SE (2001). Influences of self-beliefs, social support, and comfort in the university environment on the academic non-persistence decisions of American Indian undergraduates. *Cul Divers Ethnic Minor Psychol* 7, 88–102.
- Graham M, Frederick J, Byars-Winston A, Hunter A-B, Handelsman J (2013). Increasing persistence of college students in STEM. *Science* 341, 1455–1456.
- Hu L, Bentler P (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Model* 6, 1–55.
- Hurtado S, Cabrera NL, Lin MH, Arellano L, Espinosa LL (2009). Diversifying science: underrepresented student experiences in structured research programs. *Res High Educ* 50, 189–214.
- Hurtado S, Han JC, Sáenz VB, Espinosa LL, Cabrera NL, Cerna OS (2007). Predicting transition and adjustment to college: biomedical and behavioral science aspirants' and minority students' first year of college. *Res High Educ* 48, 841–887.
- Jackson DL, Arthur J, Purc-Stephenson R (2009). Reporting practices in confirmatory factor analysis: an overview and some recommendations. *Psychol Methods* 14, 6–23.
- Jöreskog KG, Sörbom D (1982). Recent developments in structural equation modeling. *J Marketing Res* 19, 404–416.
- Kelman HC (2006). Interests, relationships, identities: three central issues for individuals and groups in negotiating their social environment. *Annu Rev Psychol* 57, 1–26.
- Koenig R (2009). U.S. higher education. Minority retention rates in science are sore spot for most universities. *Science* 324, 1386–1387.
- Laursen S, Hunter A-B, Seymour E, Thiry H, Melton G (2010). *Undergraduate Research in the Sciences: Engaging Students in Real Science*. Hoboken, NJ: Wiley.
- Lent RW, Brown SD, Hackett G (1994). Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *J Vocat Behav* 45, 79–122.
- Lent RW, Brown SD, Larkin KC (1986). Self-efficacy in the prediction of academic performance and perceived career options. *J Couns Psychol* 33, 265–269.
- Lent RW, Brown SD, Sheu H-B, Schmidt J, Brenner BR, Gloster CS, Wilkins G, Schmidt LC, Lyons H, Treistman D (2005). Social cognitive predictors of academic interests and goals in engineering: utility for women and students at historically black universities. *J Couns Psychol* 52, 84–92.
- Lent RW, Lopez FG, Bieschke KJ (1991). Mathematics self-efficacy: sources and relation to science-based career choice. *J Couns Psychol* 38, 424–430.
- Lundy-Wagner V (2013). Is it really a man's world? Black men in science, technology, engineering, and mathematics at historically black colleges and universities. *J Negro Educ* 82, 157–168.
- Marin G (1993). Defining culturally appropriate community interventions: Hispanics as a case study. *J Community Psychol* 21, 149–161.
- McDonald RP (1989). An index of goodness-of-fit based on noncentrality. *J Classif* 6, 97–103.
- National Center for Education Statistics (2014). *Digest of Education Statistics, 2014*. [https://nces.ed.gov/programs/digest/d14/tables/dt14\\_318.45.asp](https://nces.ed.gov/programs/digest/d14/tables/dt14_318.45.asp) (accessed 6 January 2016).
- National Science Foundation (2013). *Women, Minorities, and Persons with Disabilities in Science and Engineering, Special Report NSF 15-311*. Arlington, VA.
- Netemeyer RG, Bearden WO, Sharma S (2003). *Scaling Procedures: Issues and Applications*. Thousand Oaks, CA: Sage.
- Pfund C, Pribbenow CM, Branchaw J, Lauffer SM, Handelsman J (2006). The merits of training mentors. *Science* 311, 473–474.
- 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.
- Seymour E, Hewitt N (1997). *Talking about Leaving*. Boulder, CO: Westview.
- Thew D, Smith SR, Chang C, Starr M (2012). The deaf strong hospital program: a model of diversity and inclusion training for first-year medical students. *Acad Med* 87, 1496–1500.
- Trujillo G, Tanner KD (2014). Considering the role of affect in learning: monitoring students' self-efficacy, sense of belonging, and science identity. *CBE Life Sci Educ* 13, 6–15.
- Usher EL, Pajares F (2006). Sources of academic and self-regulatory efficacy beliefs of entering middle school students. *Contemp Educ Psychol* 31, 125–141.
- Valantine HA, Collins FS (2015). National Institutes of Health addresses the science of diversity. *Proc Natl Acad Sci USA* 112, 12240–12242.
- Weston T, Laursen S (2015). The Undergraduate Research Student Self-Assessment (URSSA): validation for use in program evaluation. *CBE Life Sci Educ* 14, ar33.