An Overview on Research on Gender and Under-Represented Ethnicities in Physics Education

Laura McCullough

Department of Chemistry and Physics
University of Wisconsin-Stout
Menomonie, WI 54751
McCulloughL@uwstout.edu

Abstract:

This chapter provides an overview of the physics education research focusing on two under-represented populations in physics: women and under-represented races and ethnicities. A brief overview of the historical data on the participation of women and under-represented races leads into some of the questions that PER has asked regarding how physics education may differently affect these groups. Suggestions for those interested in doing such research are provided.

1. Introduction:

Women and some races and ethnicities have been under-represented in physics throughout its history. In the last fifty years, a great deal of research has been conducted on how to improve participation rates for these populations. While no panacea has been found, research continues to discover things that help and hinder participation by these groups. In physics education research (PER), people are starting to focus efforts on ways in which physics education and the physics classroom can promote or hamper progress on equitable participation. The goal of this paper is to provide readers with a brief foundation of the work that has been done in PER and related fields.

Nota Bene: The language that scholars use to discuss these issues is in the midst of a period of dynamic change, and terms that may have a significant body of research and outreach associated with them are being replaced by newer, more precise and inclusive language. For example, I will be using the acronym URE for under-represented races and ethnicities in some places. In others—mostly where I am talking about established committees or referring to older research—I will stick with the language choices those groups and study authors use for clarity and continuity's sake.

2. Mind the gap: Women and UREs in Physics:

The issue of low participation of women and UREs in STEM fields is significant, and of national importance.[1] Many organizations keep track of issues regarding gender and UREs. The National Science Foundation has been collating data on women and minorities in STEM fields for decades, publishing the data in a biennial report titled "Women, Minorities, and Persons with Disabilities in Science and Engineering".[2] In the field of physics, the American Institute of Physics has a regular series of publications on women and minorities in physics. The American Physical Society, the American Institute of Physics and the American Association of Physics Teachers all have committees or groups focused on women and minorities.

Low participation rates are problematic for multiple reasons. It is not just a matter of equity for equity's sake. We are losing out on talented people who have an interest in physics because the field is not welcoming to them. This talent loss is happening at a time when we as a country need to strengthen our STEM education to remain competitive internationally.[3] Diversity in a field creates a stronger field, not just ecologically but also sociologically.[4] We need to promote the participation of under-represented groups in physics both for their benefit and for the benefit of the field and society.

Most physics-related funding agencies and professional organizations view this as an issue of concern and recognize the need to do better. The low rates of women and some races and ethnicities have led to significant funding and research efforts on the topic. Figures 1 and 2 show how low these rates tend to be.

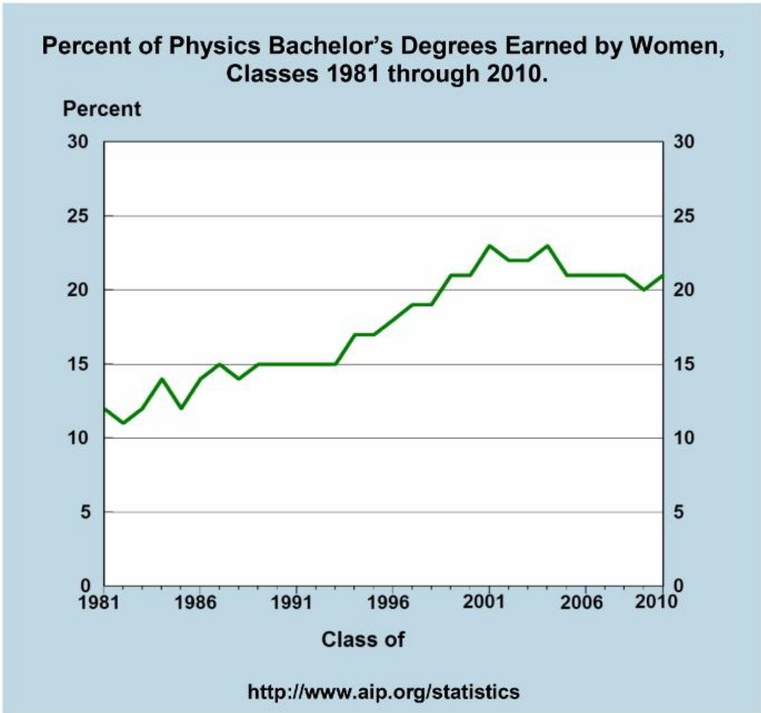

Figure 1: Percent of Physics Bachelor's Degrees Earned by Women[5]

Minority and Ethnic Profile of Physics Bachelor's, Classes of 2008 through 2010. (3-Year Average)

	Number	Percent of all Physics Bachelor's
White	4,808	81
Asian American	310	5
Hispanic American	221	4
African American	157	3
Other US citizens	70	1
Non-US citizens	333	6
Total	5,899	100%

http://www.aip.org/statistics

Figure 2: Minority Profile of Physics Bachelor's[6]

With about 80% of physics bachelor's degrees going to men, and about 80% to whites, there is a lot of room for research on why women and UREs are less likely to go into physics. The gender and science literature is large and has a long history. How does PER fit in to this broader picture?

Physics education researchers want to discover what the educational environment (formal or informal) can do to help or hinder the participation of these populations. How might a physics course in college affect a young Latina's likelihood to continue in physics? Does a test or textbook full of white male figures influence how non-white males respond to the test or text? Can student-centered, interactive pedagogies decrease the gender gap in physics? Just as PER improves physics education, PER can also improve the physics education experience for those who are historically less likely to continue in physics.

Another important reason for PER to study gender and issues of race and ethnicity is that the majority of socialization of physicists starts in the classroom.[7] Physics classes are where students often first see other physics people and learn what the field and those in it are all about and how they interact with each other and the material.

Population shifts in the U.S. mean that the proportion of the total population that is white and male is becoming steadily smaller, and that is likely to continue. This makes it increasingly important that non-whites and non-males feel welcome in our fields of physics and PER. PER has already had an impact on physics education; we can also have an impact on how physics and physics education promote the participation of under-represented groups. In order to help further that goal, this paper will look at where we started, and where we are headed.

3. Humble beginnings:

The earliest article about gender or UREs in PER might surprise you, since PER is so young itself. A 1976 article in the American Journal of Physics discussed education materials designed to support women's participation in physics.[8] I classify this as PER rather than straight gender and science research because of methodology. The authors created educational materials and tested the efficacy of those materials using a pretest/posttest questionnaire.

After this article, there is quite a gap before more gender and PER shows up on the scene. The next PER article with a nod to gender in the American Journal of Physics is in 1992, in which the authors suggest that teachers using cooperative groups for problem solving should take into account the gender of the group members.[9] UREs first appear in a 1999 article about a program for at-risk students at Rutgers University.[10]

It is in the late 1990s that research on gender and UREs begins to have real presence in the field. Since then, focus on these topics has become a part of research in most of the subfields within PER. How did we go from gender and URE research being invisible or an afterthought, to being a standard or expected part of most studies? In much the same way that PER has become such an important part of physics education: by demonstrating to people that there is an issue that needs to be addressed. As with early studies in PER, most of the research on gender in physics education started with conceptual tests.

4. Women, UREs, and conceptual tests:

The most widely known conceptual test in PER is the Force Concept Inventory (FCI).[11] Gender and race-based differences were not examined in the first articles about it and its predecessor (the Mechanics Baseline Test).[12] Two years later, the Test of Understanding Graphs-Kinematics was published and there was a brief mention of a gender difference in the initial testing: "the mean for males of 9.5 was significantly better than the 7.2 value for females."[13] The second overview test for first semester physics, the Force and Motion Conceptual Evaluation, was published in 1998 and did not include women or UREs in its discussion.[14]

For second-semester physics, the Conceptual Survey of Electricity and Magnetism did not mention gender or race in their initial publication in 2001.[15] A few years later, though, in 2004, the Determining and Interpreting Resistive Electric Circuits Concept Test (DIRECT) test authors noted "significant differences were found in the averages for males and females with males outperforming females at all levels".[16]

The two biggest tests of students' attitudes towards physics are another good example of an increasing awareness of gender in the field. The initial study (in 1997) of the Maryland Physics Expectations Survey (MPEX)[17] did not include gender, but in 2006 the Colorado Learning Attitudes about Science Survey (CLASS) trial data was analyzed to find "the responses to nearly half the statements show significant gender differences."[18]

Once the authors of a test publish their instrument, the big question is how do students perform on the test across the country, or across the globe? For the most widely-known and widely-used test, the FCI, there is a significant performance gap favoring males. One of the earliest American Association of Physics Teachers talks on gender and the FCI found that men outperform women on the pre-test, the post-test, and on relative gain.[19] This gap is present in high school classrooms and college classrooms, at Harvard and at community colleges. Over 15 years later, the gender gap on the FCI remains.[20] The latest research literature suggests that the gap is caused by a combination of small factors rather than any one major difference.

Unfortunately, there is still very little data on UREs and conceptual tests. A dissertation by Watkins[21] suggests that UREs perform more poorly on the FCI than their white counterparts, but there is little else in the extant published research. This dissertation provides important reading for those interested in these issues.

As more and more researchers in physics education have become aware of the importance of studying our classrooms and materials for their effects on women and UREs, there has been a significant increase in the number of articles looking at differences on conceptual tests. One important article for those interested in studying this is Willoughby and Metz's paper on different gain calculations for conceptual tests.[22] They do a nice job discussing different gain calculations and how different calculations may show different gender patterns. This sort of work can be a good entry point into this subfield, as many people have access to both conceptual test data and demographic data for their own classes.

5. Expanding Horizons:

From its early work on gender differences on conceptual tests, PER has expanded its focus to include looking at gender differences on online homework,[23] and initial forays into whether or not the active-learning, student-centered pedagogies we study may have differential effects by gender. In the first major study of pedagogy and gender in PER, Lorenzo, Crouch, and Mazur studied FCI gain scores for men and women in different types of courses (traditional, partly interactive, fully interactive). They concluded that fully interactive courses reduced the FCI gender gap for Harvard students.[24]

Another expansion of early work includes validating the CLASS test at a Hispanic-serving institution, where the authors find that the CLASS is a valid test for this URE population.[25] This study is counterbalanced by another very interesting article by Sabella et al. suggesting that our typical PER tools (conceptual tests, inventories) may not translate well into non-majority institutions such as primarily black, urban institutions.[26]

6. Blossoming Possibilities:

In the last decade there has been a significant increase in the numbers of people and studies examining women and UREs in PER. Not only are researchers often including measures to address gender and race when studying something else, but there are numerous studies focusing primarily on gender and physics education. Thus far, gender issues and effects are more often studied, leaving race/ethnicity an understudied area with enormous possibilities for the advancement of knowledge.

One of the more important questions being investigated is whether or not student-centered, active-engagement pedagogies can help reduce the gender gap in physics and physics education. We have known for a long time that more active pedagogies can improve student performance.[27] Early work outside PER has suggested that poor pedagogy can turn people away from physics; pedagogy was the third-highest rated reason for leaving science in a seminal work by Seymour and Hewitt.[28] Reports of poor teaching in [science, math and engineering] classes were by far the most common complaint of all switchers and non-switchers. Another study (now twenty years old) reports that science teachers are more likely to grade on a curve and are less likely to use active learning techniques.[29] While we do not yet know about specific impacts on every population, we can be glad for most students that high school teaching has come a long way toward adopting better pedagogy since then![30]

In PER, there are mixed reports on the effectiveness of active-learning techniques on reducing the gender gap. Pollock et al. found results that did not support the Harvard results: interactive engagement did not help reduce gender differences.[31] In later studies, this group concludes that it is an accumulation of small differences that add up to significant differences in results by gender in physics courses.[32]

Several studies on UREs have looked at how active-engagement pedagogies may help minority populations. The Student-Centered Active Learning Environment with Upside-down Pedagogy (SCALE-UP) classrooms have been shown to reduce failure rates in female and minority students (Beichner and Saul, 2003).[33] Similarly, a program at Rutgers University has reduced failure rates for at-risk populations, many of whom are UREs.[34]

One gender-focused article in PER has been featured in the pages of *Science*,[35] examining a short in-class intervention on values affirmation that helped reduce the gender gap in an introductory physics course.

The expansion of the field means that there are far more gender- and race-related articles being published than can be mentioned here, but a good deal of work has come out of University of Colorado-Boulder, and Florida International University. The new research has pushed the boundaries beyond performance gaps and curricular issues, looking into bias, attitudes, affirmation and sense of belonging, and other people-focused issues. Doing a search for gender or race on the PRST-PER website[36] provides wonderful reading material. The prevalence of gender research in PER has resulted in a focused collection on gender in physics, published in the summer of 2016.[37]

7. Getting Started:

For those readers interested in doing work related to women and UREs, there is important research to read, and important ideas to keep in mind.

7.1 Stereotype threat

Anyone doing work in these areas should familiarize themselves with the concept of stereotype threat. While there are many articles in many areas on this, a great introduction is Claude Steele's book "Whistling Vivaldi".[38] Stereotype threat is the idea that our actions are changed by the awareness, or even the presence, of stereotypes about us. White men told that they are going to be tested on basketball performance as compared to black men perform more poorly, unconsciously doing worse because of the stereotype that black men are better basketball players. When told that a math test tends to discriminate by gender, women perform worse and men perform better. This usually happens unconsciously, but the effect is very well documented.[39] A good educational resource with a strong bibliography is the website http://www.reducingstereotypethreat.org/.

One of the most important things to come out of stereotype threat research for PER is that all demographic questions should be placed at the end of a questionnaire/test/survey rather than at the beginning. Even the presence of a M/F-type question at the beginning of a test can trigger stereotype threat.[40]

The idea of implicit bias is a related concept that is further important reading for those getting started.[41] Harvard's Implicit Association Test[42] is a quick and easy introduction to the idea, and also allows a researcher to discover her or his own hidden biases.

7.2 The idea of significance and statistics

One of the issues that is important when looking at gender or race differences in PER studies is the idea of statistical significance versus educational significance. If twenty studies were to say that men outperform women on a test—but not statistically significantly—it does not mean that there is no gender gap. Likewise, if a study with 2000 people shows a statistically significant racial gap of 1 question out of 75, the educational significance of the study is minimal, though real.

Statistics can hide important differences and trends. And, as mentioned above, how you analyze your data may include hidden biases. Researchers still need to study how different calculations or methodologies may influence examinations of gender and racial differences, and then we need to make our field aware of these confounds. Another important resource regarding this area is an article by Rodriguez et al.[43] discussing models of equity and how they affect analysis results. They reviewed the aforementioned Harvard data under different models and note different interpretations that could be made.

A common mistake when people first start examining gender or URE effects in data is to take all their data and run statistical tests on every semester, every class, to see if there are gaps. Unfortunately, the nature of statistics means that at a p-value of 0.05, if you run 20 tests, you are likely to get one that is statistically significant, just due to random chance. Every time you run another test, you are increasing the chances of getting a "false positive". So researchers should be thoughtful about what tests to run. Develop hypotheses to test first; figure out where you believe there may be differences or effects. Then determine the best test to run to test that hypothesis.

A better way than a simple t-test to determine if a statistical difference is meaningful is to calculate the effect size.[44] There is also a body of work on Differential Item Functioning (DIF) for multiple-choice type tests.[45]

7.3 Different Approaches

Another important point for incoming researchers to consider in their focus is their motivation for studying gender and race or ethnicity issues. Some people may approach it from the perspective of finding an immediate solution to a problem. Others might be more interested in a thorough understanding of the problem and its underlying causes. Your motivation will lead you to approach and structure research differently.

A purely quantitative approach will give you different information than qualitative approaches. Many physicists are most comfortable with numbers and statistics, but as we learn more about gender and culture, it is clear that our attitudes and social settings have a large effect on performance. Consider what you want to know, and learn about the best sort of data to gather. Social theory and psychology have much to offer both researchers and teachers. Qualitative work often involves interviewing appropriate groups and coding results, grouping into themes, and developing theories from the patterns that emerge.[46] There are numerous types of qualitative research and many resources to help someone move into this area.

7.4 Biases

As has been mentioned previously, we all come into our research with inherent biases, many of which we are unaware of. The biases we have will affect how we choose our data, our analyses, our conclusions. The best counter to this is an awareness of those biases (see the Implicit Association Test section above), and education about how different analyses can provide different results (see the articles above by Rodriguez and Willoughby). There is a common belief (explicit or inherent) that white men are the standard to which all others are, and should be, judged. This "deficit model" can cause bias in research as well as teaching: it can lead us to ask how to get women to perform like men, when a better question is to examine what causes the differences between groups. An example: If women don't do well on a test designed by men and tested on men, administered in a classroom full of men, is the problem really with the women?

7.5 Flexibility

One of the advantages of research on gender and race issues is that it is possible to dabble around the edges or dive in completely. There is so much ground to cover in our need for more understanding of these issues that everyone should be encouraged to look at gender and racial/ethnicity effects in their research at whatever level they can accommodate. It need not be the central focus of all research, but it is simple enough to add some demographic questions at the end of a survey or test in order to get additional research data.

If you are interested in gender or race effects in your own classroom, you can dive right in. This sort of work may not result in article-worthy research, but can provide useful information to the individual. If you want to conduct research, you should take care to learn about appropriate methods and analyses, just as with any subfield, especially if you are hoping to publish.

Another advantage to this field is that there are numerous possibilities for disseminating work and for finding it. Beyond the major PER publishers, there are science teaching journals such as the *Journal of College Science Teaching*, *Physics Education*, and the *Journal of Research on Science Teaching*. Focusing on UREs and women, there are also the *Journal of Women and Minorities in Science and Engineering, the Journal of Multicultural, Gender and Minority Studies* and the *Journal of Research in Gender Studies*. There is an enormous amount of research outside PER that can help a person within PER learn about studying under-represented groups.

7.6 A Point of Concern

One possible issue for those interested in conducting research in gender and URE areas to be aware of is that this subfield can be somewhat marginalized in certain settings. Bringing up race/ethnicity and gender sometimes makes people uncomfortable, and this could be an issue for the researcher if those uncomfortable people are a search committee, or a boss, or an advisor. This is unfortunate and should not be the case, but the researcher interested in this area should be aware of possible personal or institutional biases that may play against them.

8. Future Directions:

As the research base concerning gender and race/ethnicity in PER grows, this area continues to need people actively working on the basic question of how the physics classroom affects under-represented races/ethnicities and women. If we are to continue to promote physics participation for all, we need to continue studying how physics education helps or hinders these efforts. I expect that in the future, our definition of gender will become broader, just as “Women’s Studies” has grown into “Gender Studies”. I also expect to see a similar expansion of studies focused on UREs. We should widen our view of “under-represented” groups to include persons with disabilities and members of the LGBTQ community. To do this, we need to make sure that people in PER are aware of the importance of studying these topics.[47]

9. Conclusions:

As the need for more scientists and technically trained people continues to grow, and the need for strong science education becomes more desperate, it behooves the PER community to broaden the field to include the effects of physics education on those who do not match the traditional picture of a physicist. Research on under-represented groups is an important part of PER that everyone should be aware of. We still have many questions that need to be answered, and newcomers have excellent opportunities to make contributions. Step up and join us in making a difference!

---

[1] National Research Council, *Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future.* (The National Academies Press, Washington DC, 2007).

National Research Council, *Rising Above the Gathering Storm, Revisited: Rapidly Approaching Category 5.* (The National Academies Press, Washington, DC, 2010).

[2] National Science Foundation, *Women, Minorities, and Persons with Disabilities in Science and Engineering.* NSF Report 13-304. (2013). Available at http://www.nsf.gov/statistics/wmpd/2013/start.cfm

[3] http://www.jec.senate.gov/public/index.cfm/democrats/2012/4/stem-education-preparing-jobs-of-the-future

http://www.ed.gov/stem

[4] W. W. Wittmann, "Group Differences in Intelligence and Related Measures" in *Understanding and Measuring Intelligence* ed. Wilhelm O and Engle R (Sage, London, 2005).

[5] American Institute of Physics, Figure: *Percent of Physics Bachelor's Degrees Earned by Women.* http://www.aip.org/statistics/trends/highlite/edphysund/figure6a.htm

[6] American Institute of Physics, Figure: *Minority and Ethnic Profile of Physics Bachelor's.* http://www.aip.org/statistics/trends/highlite/edphysund/table1e.htm

[7] B. Whitten, S. Foster, M. Duncombe, H. Zorn, "'Like a Family': What Works to Create Friendly and Respectful Student-Faculty Interactions", J. Women Min. Sci. Eng.**10**, 3. (2004).

B. Whitten, S. Foster, M. Duncombe, P. Allen, P. Heron, L. McCullough, K. Shaw, B. Taylor, H. Zorn. "What Works? Increasing the Participation of Women in Undergraduate Physics", J. Women Min. Sci. Eng. **9**, 3-4. (2003).

[8] D. L. Moché, "Development of Educational Materials to Recruit Women into Scientific Careers", Am. J. Phys. **44**, 390 (1976).

[9] P. Heller and M. Hollabaugh, "Teaching Problem Solving Through Cooperative Grouping. Part 2: Designing Problems and Structuring Groups", Am. J. Phys. **60**, 637 (1992).

[10] E. Etkina, K. Gibbons, B. L. Holton, and G. K. Horton, "Lessons Learned: A Case Study of an Integrated Way of Teaching Introductory Physics to at-risk Students at Rutgers University", Am. J. Phys. **67**, 810 (1999).

[11] D. Hestenes, M. Wells, and G. Swackhamer, "Force Concept Inventory," The Physics Teacher, **30**, 141-158. (1992).

[12] D. Hestenes and M. Wells, "A Mechanics Baseline Test," The Physics Teacher, **30**, 159-165 (1992).

[13] R. Beichner, "Testing Student Interpretation of Kinematics Graphs," Am. J. Phys., **62**, 750-762, (1994), pg 752.

[14] R. Thornton, and D. Sokoloff, "Assessing Student Learning of Newton's laws: The Force and Motion Conceptual Evaluation and the Evaluation of Active Learning Laboratory and Lecture Curricula", Am. J. Phys., **66**, 338-352 (1998).

[15] D. Maloney, T. O'Kuma, C. Hieggelke, and A. Van Heuvelen, "Surveying Students' Conceptual Knowledge of Electricity and Magnetism", Am. J. Phys. **69**, S12 (2001).

[16] P. Engelhardt, and R. Beichner, "Students' Understanding of Direct Current Resistive Electrical Circuits", Am. J. Phys, **72**, 1 (2004), p. 105.

[17] E. F. Redish, R. N. Steinberg, and J. M. Saul, "Student Expectations in Introductory Physics," Am. J. Phys. **66**, 212-224 (1998).

[18] W. K. Adams, K. K. Perkins, N. S. Podolefsky, M. Dubson, N. D. Finkelstein, and C. E. Wieman, "New Instrument for Measuring Student Beliefs about Physics and Learning Physics: The Colorado Learning Attitudes about Science Survey", Phys. Rev. ST Phys. Educ. Res. **2**, 010101 (2006), p. 9 of 14.

[19] L. McCullough, "Gender Differences on Multiple-Choice Conceptual Tests" In *Winter Meeting of the American Association of Physics Teachers* (Phoenix, AZ, 1997).

[20] L. McCullough, and C.H. Crouch, "Gender, Educational Reform, and Instructional Assessment, Part 1." In *Summer Meeting of the American Association of Physics Teachers.* (Philadelphia, PA, 2002). http://www.uwstout.edu/faculty/mcculloughl/upload/McCulloughAAPTJan02.pdf;
J. Docktor and K. Heller, "Gender Differences in Both Force Concept Inventory and Introductory Physics Performance," *2008 Physics Education Research Conference* (AIP, 2009). http://perusersguide.org/items/detail.cfm?ID=7958;
R. D. Dietz, M. R. Semak, C. W. Willis, "The Gender Gap on the FCI—Question by Question," in *Winter Meeting of the American Association of Physics Teachers*

(Washington, DC, 2010)
http://www.unco.edu/nhs/physics/files/AAPTW09Poster.pdf;
V. P. Coletta, J. Phillips, and J. J. Steinert, "FCI Normalized Gain, Scientific Reasoning Ability, Thinking in Physics, and Gender Effects," *2011 Physics Education Research Conference* (AIP, 2012).
http://www.compadre.org/per/items/detail.cfm?ID=11803
A. Madsen, S. McKagan, and E. Sayre, "Gender Gap on Concept Inventories in Physics: What is Consistent, What is Inconsistent, and What Factors Influence the Gap?" Phys. Rev. ST Phys. Educ. Res. **9**, 020121 (2013).

[21] J. Watkins, *Examining issues of underrepresented minority students in introductory physics.* Ph.D. Thesis, Harvard University, 289 pages, (2010).

[22] S. Willoughby and A. Metz, "Exploring Gender Differences with Different Gain Calculations in Astronomy and Biology," Am. J. Phys. **77**, 651 (2009).

[23] G. Kortemeyer, "Gender Differences in the Use of an Online Homework System in an Introductory Physics Course," Phys. Rev. ST Phys. Educ. Res. **5**, 010107 (2009).

[24] M. Lorenzo, C. H. Crouch, and E. Mazur, "Reducing the Gender Gap in the Physics Classroom," Am. J. Phys. **74**, 118 (2006).

[25] V. Sawtelle, E. Brewe, and L. Kramer, "Validation Study of the Colorado Learning Attitudes about Science Survey at a Hispanic-serving Institution," Phys. Rev. ST Phys. Educ. Res. **5**, 023101 (2009).

[26] M. Sabella, K. Coble, and S. P. Bowen, "Using the Resources of the Student at the Urban, Comprehensive University to Develop an Effective Instructional Environment," Physics Education Research Conference 2008: Edmonton, Canada. Volume 1064, (2008).

[27] R. Hake, "Interactive-engagement versus Traditional Methods: A Six-Thousand-Student Survey of Mechanics Test Data for Introductory Physics Courses," Am. J. Phys. **66**, 1 (1998).

[28] E. Seymour and N. Hewitt, *Talking About Leaving: Why Undergraduates Leave The Sciences.* (Westview Press, Boulder, CO, 1997).

[29] J. Milem and H. Astin, "Scientists as Teachers: A Look at their Culture, Their Roles, and Their Pedagogy," Paper presented at the NARST meeting, New Orleans, LA. (1994).

[30] Z. Hazari, R. H. Tai, and P. M. Sadler, "Differences in Introductory University Physics Performance: The Influence of H.S. Physics Preparation and Affective Factors," Science Education, **91**, 847-876 (2007).

[31] S. J. Pollock, N. D. Finkelstein, and L. E. Kost, "Reducing the Gender Gap in the Physics Classroom: How Sufficient is Interactive Engagement?" Phys. Rev. ST Phys. Educ. Res. **3**, 010107 (2007).

---

[32] L. E. Kost-Smith, S. J. Pollock, and N. D. Finkelstein, "Gender Disparities in Second-Semester College Physics: The Incremental Effects of a 'Smog of Bias'," Phys. Rev. ST Phys. Educ. Res. **6**, 020112 (2010)
L. E. Kost, S. J. Pollock, and N. D. Finkelstein, "Characterizing the Gender Gap in Introductory Physics," Phys. Rev. ST Phys. Educ. Res. **5**, 010101 (2009).
[33] R. Beichner, and J. Saul, "Introduction to the SCALE-UP (Student-Centered Activities for Large Enrollment Undergraduate Programs) project," *Proceedings of the International School of Physics* "Enrico Fermi," Varenna, Italy (2003). Available at http://www.ncsu.edu/PER/Articles/Varenna_SCALEUP_Paper.pdf.
[34] S. Brahmia, "Improving Learning for Underrepresented Groups in Physics for Engineering Majors," Physics Education Research Conference 2008: Edmonton, Canada. Volume 1064, Pages 7-10 (2008).
[35] A. Mikaye, L. Kost, N.D. Finkelstein, S. Pollock, G. Cohen, T. Ito, "Reducing the Gender Achievement Gap in College Science: A Classroom Study of Values Affirmation," Science **26**, Vol. 330 no. 6008 (2010).
[36] https://journals.aps.org/prstper/
[37] https://journals.aps.org/prper/edannounce/10.1103/PhysRevPhysEducRes.12.020001
[38] C. Steele, *Whistling Vivaldi: How Stereotypes Affect Us and What We Can Do.* (W. W. Norton & Co, 2011).
[39] C. Steele and J. Aronson, "Stereotype Threat and the Intellectual Test Performance of African Americans," J. Pers. Soc. Psych., **69,** 797–811 (1995).
S. Stroessner, and C. Good, "Stereotype Threat: An Overview," http://diversity.arizona.edu/sites/diversity/files/stereotype_threat_overview.pdf
[40] K. Danaher and C. Crandall, "Stereotype Threat in Applied Settings Re-Examined," J. Appl. Soc. Psych. **38**, 6 (2008).
[41] C. Staats and C. Patton, *State of the Science: Implicit Bias Review.* (2013.) http://kirwaninstitute.osu.edu/docs/SOTS-Implicit_Bias.pdf
[42] IAT available at https://implicit.harvard.edu/implicit/
[43] I. Rodriguez, E. Brewe, V. Sawtelle, and L. H. Kramer, "Impact of equity models and statistical measures on interpretations of educational reform," Phys. Rev. ST Phys. Educ. Res. **8**, 020103 (2012).
[44] R. Coe, "It's the Effect Size, Stupid: What Effect Size is and Why it is Important," Paper presented at the Annual Conference of the British Educational Research Association, September 2002. http://www.leeds.ac.uk/educol/documents/00002182.htm.
[45] A. Schmitt, P. Holland, and N. Dorans, "Evaluating Hypotheses About Differential Item Functioning," Paper presented at Educational Testing Service/AFHRL Conference, Oct 6 1989. http://files.eric.ed.gov/fulltext/ED387525.pdf.

---

[46] Association for Educational Communications and Technology, "Introduction to Qualitative Research", http://www.aect.org/edtech/ed1/40/40-01.html.

[47] A. Traxler, X. Cid, J. Blue, and R. Barthelemy, "Enriching gender in physics education research: A binary past and a complex future," Phys. Rev. Phys. Educ. Res. **12**, 020114 (2016).