Book chapter: To render the extraordinary, ordinary: acknowledging bias and barriers

A transcript of this talk, invited to be included in the 2018 ACS refereed volume, as given is currently being edited. However, here are notes written prior to talk. -October, 2018

About mentoring:

- Having students affirm their skills and values
  e.g. Undergraduate writing exercises, Miyake et al (2010)
- Women teachers and role models
  e.g. Air Force Academy study, Carrell et al (2009)
- Having men on board (enlist the alpha males!)
  e.g. … girls' interest in math decreases as their fathers' gender stereotypes increase,
  whereas boys' interest in math increases as their fathers' gender stereotypes
  increases, *P. Davis-Kean et al (2007)*
- Mentoring … vertical and horizontal
  e.g. Women with mentor are much (93% vs. 68%) more likely to receive grants, Rachel Ivie, AIP data (2010) and engineers with sponsor are 3 times more likely to see ideas implemented (Laura Sherbin, Columbia U)

The way in which women are regarded by their students might be both a cause and an effect of the fact that many fewer women than men establish lifelong careers in science, technology, engineering, mathematics (STEM) fields (Ceci and Williams, 2011), occupying only 9% to 16% of tenure-track positions (Nelson and Brammer, 2007). It is beneficial for minority

students when the race or gender of student and professor resonate (Blake-Beard et al., 2011), as the professor as role model has been shown to be crucial in retaining underrepresented students (e.g. Carlone and Johnson, 2007; Hughes, 2012; Kosoko-Lasaki et al., 2006). Retaining women professors yields greater numbers of women students (e.g. Drury et al., 2011). One might expect this for many reasons. For one, psychological studies (e.g. Murphy et al., 2007) indicate that women feel a greater sense of belonging in a STEM setting where the population is gender- balanced. For another, the types of psychosocial support that were needed, as for example by STEM student women of color in Johnson's (2007) survey study, are found to be provided more frequently by ingroup mentors (e.g. Downing et al., 2005; Manke and Cohen, 2011).

When a potential role model leaves academia, the loss has a domino effect, harming the careers of potential students and coworkers. While real-world survey data (Downing et al., 2005; Hazari et al., 2013) show that role models need not be women in order to encourage interest in women to pursue STEM, once in the discipline, identification with a female role model "inoculates" one against stereotype threat (Manke and Cohen, 2011; Marx and Roman, 2002; Stout et al., 2011). Self-similar role models can not only improve underrepresented students' performance in STEM courses, but are significantly more likely to enhance their self-concept and retain them in the discipline (Drury et al., 2011; Newman, 2011; Stout et al., 2011).

Better representation of all types of people is not only important from a social-justice point of view but also having a diversephysics community is Step 7 of Whitten's (2012) "(Baby) steps toward a feminist physics." And it is what Harding (2001) would consider a marker of the quality of the scientific knowledge which a community can generate.

A great deal of research on women in STEM fields suggests that a number of socio-cultural and psychological factors are

to blame for lesser numbers and lower job satisfaction (Carrell et al., 2009; Ceci and Williams, 2011; Steele et al., 2002; Valian, 1998). One of these factors, the motivation for the current study, is the evidence of gender-biased evaluations. Negatively biased evaluations harm careers in a variety of ways: from low self-concept and low job satisfaction to inequities in hiring and promotion. The importance of recognition as a capable "science person" was shown to influence persistence in research on women of color by Carlone and Johnson (2007). There is an important climate issue of being a good "fit" within an institution (Gallagher and Trower, 2009; Yost et al., 2013).

There is a substantial body of research on the role of gender and student evaluations. Student evaluations are often crucial in decisions to retain and promote faculty (e.g. Benton and Cashin, 2014). Yet ratings show that teachers are held accountable to gendered expectations, with a larger and often inconsistent set of expectations placed on women (Anderson and Smith, 2005; Basow, 1998; Bennet, 1982; Sinclair and Kunda, 2000). Men are sometimes rated as more effective (Reid, 2010), although it is essential to delineate different dimensions of teaching effectiveness (e.g. knowledge, organization, approachability, enthusiasm, etc.) as well as details like whether the course is a large lecture (female unfriendly) or small discussion (Martin, 2013). Though ethnicity was not a variable in our study, it is known that both the race of the professor and interactions between race and gender of the professor (e.g. Ho et al., 2009; Reid, 2010) and/or the student (e.g. Anderson and Smith, 2005; Sinclair and Kunda, 2000; Sprinkle, 2008) can be important on evaluations. Evaluation of professors in STEM fields, physics in particular, are less well-studied in isolation due to the difficulty of obtaining good statistics from small numbers of students (Basow, 1995, 1998, 2011). But, for example, in Potvin et al.'s (2009) study of evaluations of high school biology, chemistry, and physics

teachers, male students gave lower ratings to women teachers in all three fields, while female students gave lower ratings only to the women physics teachers.

Despite the fact that (in most countries of the world) legally-sanctioned bias in the workplace does not exist, individuals in "stereotype-incongruent" occupations (e.g. female physicist or male kindergarten teacher) can still feel the effects of implicit bias, for which experimental evidence is extensive (Brescoll et al., 2010; Greenwald and Krieger, 2006; Institute of Medicine et al., 2007; Lemm and Banaji, 1999; Valian, 1998). In a recent study, both male and female scientists were more likely to hire and provide mentorship to a fictive candidate for the job of scientific lab coordinator, if the name on the candidate's resume was male (Moss-Racusin et al., 2012). The male candidate was felt to have greater competence, and this study implicated implicit gender bias as the reason. Similarly, Reuben et al. (2014) showed that subjects acting as "hiring managers" are implicitly biased to hire men over women for a mathematical task (addition of numbers – a task at which neither men nor women have been shown to hold a gender-based advantage). Even when "managers" were given the "applicants'" actual scores on a trial run, while the gender imbalance was mitigated somewhat, they tended to choose lower scoring men over higher scoring women.

Davison and Burke (2000) conducted a meta-analysis of nearly 50 studies that employed a simulated hiring decision using identical resumes on applications with either a female or male name. They found that when job-relevant information such as educational background and prior work experience was not provided, female job applicants faced discrimination. Studies of letters of real-world recommendation for academic and medical positions (Madera et al., 2009; Trix and Psenka, 2003; Watson,

1998) found that women received shorter letters, fainter praise and less credit for intellectual prowess and "agency" (i.e. an active, not merely reactive, researcher, and the producer of novel ideas and results). Using fictitious case-studies of job performance, Brescoll et al. (2010) found an interesting distinction: when women and men were reported to perform equally well in gender-stereotype-incongruent professions (e.g. female police chief or male president of a women's college), evaluators rated them equivalently. However, if the women and men were reported to make a mistake, those who had gender-incongruent jobs received a much harsher competence rating than their gender- congruent counterparts. This produces a double-penalty for female job-seekers, if, as in the study by Trix and Psenka (2003), more "doubt-raisers" appear in letters written for women. While both women and men in gender-incongruent professions are penalized preferentially in evaluations for making a mistake in job performance (Brescoll et al., 2010), women are uniquely punished for certain behaviors necessary in the professional world, such as giving negative feedback (Sinclair and Kunda, 2000), displaying agentic traits, or being perceived as less-than-agreeable (Heilman et al., 2004; Rudman and Glick, 1999). These results are particularly troubling, for both agency and firmness are required for a high-level position, whether in management or in academia, and particularly in STEM.

In summary, the research reviewed above suggests that female instructors in physics might be excruciatingly vulnerable to discriminatory evaluation, and disapproval of students. This male stereotypical activity is a particularly tough challenge for female instructors who must "walk the line" of being sufficiently excellent and authoritative, yet extremely nurturing to meet gender- role-based expectations of pliancy and compassion (Basow, 1995; MIT, 2011). In the current study, we investigate how well women walked this line in physics. Our testing instrument was a digitally videotaped lecture – a simulated classroom situation, with actors portraying physics professors. We

measured how students perceived one of these “college professors,” who delivered an identical lecture on an introductory physics topic. The digitally prerecorded lecture had technical content appropriate for physics students, chalkboard calculations, and a demonstration involving a laser. The scripted lecture had a built in “mistake” which the professor noticed and corrected, and the professor provided answers to a couple of questions from the class (all scripted). In these and other ways, the scripted, video-recorded lectures attempted to touch upon various performance-related issues that had proven to be gender-biased in the literature on evaluations of STEM professionals. In our study described below, each female or male student evaluated the lecture and judged the professor whom they saw for their knowledge, competency, and hireability.