

Why So Few? How to Increase the Number of Women in Science

by Meg Urry

Why Worry?

Everyone agrees there are too few women and minorities in science. But then opinions diverge, at least among scientists. Many believe that increasing diversity is a matter of social engineering, done for the greater good of society, but requiring a lowering standards and thus conflicting with excellence. Among this group are very well-meaning people who genuinely wish to increase the number of women colleagues. Yet they may be doing more harm than good.

Others understand that there are deep reasons for the dearth of women, which lead to extra obstacles to their success. Once one understands the bias against women in male-dominated fields (which has been substantiated in thousands of research studies, though usually in a literature that few natural scientists read), one must conclude that diversity in fact enhances excellence. In other words, the playing field is not level, so we have been dipping more deeply into the pool of men than of women, and thus have been unknowingly lowering our standards. Returning to a level playing field (compensating for bias) will therefore raise standards and improve our field. Diversity and excellence are aligned.

What Data Show

There are many studies documenting the differential progress of women. Long (2001) reviewed the gender dependence of salary, rank and tenure in science and engineering, using NSF data for a synthetic cohort (correcting for time since degree, type of institution, specialty, and family status). Women lag behind, in advancing and in getting tenure (see earlier similar studies by Sonnert and Holton in the 1990s). Having children has the effect of removing women from the full-time workforce, but differences for women who remain full-time are minimal (see Mason and Goulden, 2002, *Academe*, "Do Babies Matter?").

In a study of U.S. professionals in internationally-oriented business, Egan & Bendick (1994) studied how 17 factors - such as type of degree, years of experience, number of hours worked, etc. - affected the salaries of men and

women very differently. Fourteen of the 17 factors helped men more than women. For example, having a BA degree added \$28,000 on average to a man's salary but only \$9,000 to a woman's. Not constraining one's career because of a spouse added \$21,900 to the average male salary but only \$1,700 to women. Being on the "fast track" added \$10,900 for men, \$200 for women.

Some factors that enhanced men's salaries actually subtracted from women's. For example, living outside the U.S. added \$9,200 to a man's salary, on average, but subtracted \$7,700 from a woman's. Speaking a second language added \$2,600 for men and subtracted \$5,500 for women. Deliberately choosing international work added \$5,300 for men and subtracted \$4,400 for women.

Two factors helped women's salaries more than men's: negotiating for one's salary subtracted \$5,600 from men's salaries and added \$3,500 to women's. Traveling for more than 10 days per year added \$3,200 to men's salaries and \$6,300 to women's.

In a study of academic medicine, Tesch et al. (1995) showed that newly hired men get more lab and office space, more funding and more research time than women. A well-known study at MIT (1999) showed the same disparities for women and men faculty in the School of Science.

In hundreds of studies across many fields, using many measures, the advancement of women lags that of men with the same qualifications.

Why Are Women Scarce in Science?

Some of my colleagues believe women are simply not interested in science - at least, not in the physical sciences - and they do not seem to worry about the loss of talent. That is, if women are not interested, they must not be any good. Yet Xie & Shauman (2003) showed that interest in the sciences does not correlate with ability. Furthermore, they found that sex disparities in productivity (e.g., publication rates) were decreasing, and that productivity is independent of family status. Again, childbirth has the effect of removing women from full-time work, to the long-term detriment of their careers.

It is certainly true that too few high quality childcare options are available, and that women do more family care than men do. But women without children

still do not advance at the rate men do. And countries with excellent maternity and childcare benefits (e.g., Nordic countries) have some of the lowest participation of women in Physics. And finally, women with families do participate in extremely demanding careers in other scientific fields (e.g., medicine).

If it is not ability or interest, what is it? There is plenty of evidence that the playing field is not level for women and men. In 1997 Wenneras and Wold published a study in *Nature* about applications for a prestigious Swedish postdoctoral fellowship in medicine. They showed that although 46% of the applications were from women, only 20% of the fellowships were awarded to women. Reviewers of the proposals consistently gave women lower scores for the same level of productivity, and women applicants had to be 2.5 times better than men to succeed. An earlier study of peer review (Paludi and Bauer 1983) showed that papers were rated lower if the author's name was female than if it was male (initials were rated nearly as low as the female name, and subsequent interviews suggested initials were taken as a hidden indication of a female author). A recent study showed that the fraction of papers having a woman as first author increased significantly when a biology journal went to double-blind refereeing¹ (Budden et al. 2008). Studies of prizes or honors show that men receive a disproportionate number, even when one corrects for pipeline issues (Astronomy, Physics, Psychology).

There is much talk lately about “innate ability” --- perhaps women are simply not as good at science as men? This suggestion is contradicted by almost all available evidence. First of all, gender gaps in performance (for example, on math exams) are decreasing in the U.S.; if they were due to physiology, they shouldn't change so dramatically on time scales of decades. Moreover, gender gaps vary enormously by country, arguing against a genetic origin. Japanese women score better in math than U.S. men.

At the same time, gender gaps *can* be explained by culture. Research into “stereotype threat” shows that culture affects test results. A class is told they will be given a difficult math test. Men do poorly, scoring 25 of a possible 100, and women do worse, with an average grade of 10. This is the kind of gender gap that makes it into the *New York Times*: that at the extremes of performance, men substantially outscore women. However, another class is told the same story about a difficult math test, with the added information that

¹ Single-blind refereeing is when the referee knows the identity of the author but the author does not know the identity of the referee. Double-blind refereeing is when neither knows the identity of the other.

the test has been designed to be “gender neutral.” Now the women's score doubles, to 20. Interestingly, the men's score decreases, to 20. In other words, men and women score the same. These tests have been repeated many times with the same results, and have also been done to probe other stereotypes (e.g., black students perform less well than white students, when in a stereotype-threat situation, regardless of educational or socio-economic background). When the stereotype threat is activated, people under stress conform to it.

“Gender Schemas”

We are a biased society. There is no getting away from it. It is not overt: most of us think we are — and try hard to be — unbiased. It is also not men discriminating against women, it is all of us discriminating against women (and minorities). Try taking the online “implicit bias” test of Mahzarin Banaji (implicit.harvard.edu) — it is a real education. In her book “Why So Slow? The Advancement of Women,” Virginia Valian describes the origin of this bias with “gender schemas” — namely, a set of expectations of women and of men, embedded in our culture, that influence how women and men are judged.

Here are some examples of research on gender bias:

- *Heights of men and women (Biernat, Manis & Nelson 1991)* - Subjects are asked to estimate an objective quantity, namely the heights of men and women in photographs (which include some physical object like a doorway or desk to offer scale). Even though the subjects were chosen so that each gender has the same height distribution, the average height estimated for men is greater than the average height estimated for the women. We expect men to be taller - we are sure this is true (indeed, it is true at present in our society as a whole) - and so this is what we measure, even when it is not true in the particular data set.
- *Leader at table (Porter & Geis 1981)* - Undergraduate students are shown photographs of people sitting around a table, and asked to identify the leader. Where all the people pictured are men, the leader is nearly always identified as the person at the head of the table. The same is true when only women are pictured. When both men and women are pictured and a man sits at the head, he is identified as the leader. However, in the mixed gender case with a woman at the head, half the time a random man is identified as the leader.

- *Leaders talking (Butler & Geis 1990)* - Undergraduate “subjects” are shown a film of male or female students leading a discussion; the subjects are observed during the film and are asked questions about it afterward. The men in the film generate more positive facial reactions when speaking than the women, unless the women have been validated as a leader prior to the talk (e.g., with a thorough introduction covering her qualifications).
- *Establishing power through eye gaze (Dovidio et al. 1988)* - First the experimenters established that in a conversation between a superior and a subordinate (same gender), the superior looks at the subordinate while talking, but looks away when listening. The subordinate spends roughly equal amounts of time looking and listening, regardless of who is speaking. Then the experimenters showed that in conversations between men and women, men look while talking and women look while listening. This reinforces the assumption that the man is more powerful than the woman. (Note to girls: make eye contact while talking; not sure whether to look away while listening, though.)
- *Rating managers (Heilman et al. 2004)* - In a synthetic exercise where subjects are asked to rate two assistant vice-presidents in a fictitious (but heavily documented) aircraft company (a “male” environment), men are rated higher than women (despite randomized resumes) but both are deemed likeable. In a second experiment, in which women are validated prior to the evaluation (e.g., “both managers have been rated outstanding”), then men and women are rated equally competent but the woman is not likeable and is judged hostile or difficult. That is, women can be competent or likeable but not both. (Anyone notice any parallels to presidential politics?)
- *Rating resumes for a “male” job (Norton, Vandello & Darley 2004)* - Subjects are asked to rate 5 job applicants for a job in construction, based on resumes. By design, only 2 are really competent; one of the two has more education (an advanced degree in engineering and a certification from a construction industry group), and the other has more work experience (9 years compared to 5 years). In one experimental condition, the resumes are labeled with initials only; in another, the resumes are labeled names of both genders.
 - If initials, then education was judged more important than experience, and most highly educated person was ranked highest.
 - If man’s name on resume with more education, he is ranked number one.

- If woman's name is on the “educated” resume, the “more experienced” man more likely to be ranked highest and experience is described as more important in making the decision.
- *Mismatched credentials for gender-identified jobs (Uhlmann & Cohen 2005)*- Subjects fill out a questionnaire asking about the most important criteria for a gender-identified position, either a police chief (“male”) or a nursing supervisor (“female”). For example, a masculine job like police chief generally elicits more emphasis on presumptively male characteristics like physical strength, authoritative voice, and experience in law enforcement, rather than female characteristics (nurturing, feeling) such as caring or has a family. The subjects then rate applicants according to resumes that have predominately (by stereotype) “male” or “female” characteristics.
 - When a man’s name is on the resume with the male characteristics, he is ranked highest for the job of police chief.
 - However, when the woman’s name is on the resume with the male characteristics, the man is still ranked highest. In other words, the criteria change in response to the gender of the applicants.
 - Interestingly, the subjects who identified themselves in the initial questionnaire as “objective” were far more likely to change criteria (i.e., act according to gender schemas) than those who labeled themselves “not objective.” So, when someone tells you they are objective, beware.
 - When the same experiment was carried out for the stereotypically female job of nursing supervisor, the results were similar. That is, the woman was ranked highest for the job regardless of whether her qualifications aligned with those deemed most important in the initial questionnaire.
- *Sanbonmatsu, Akimoto & Gibson 1994*: 4 students pass, 4 fail, in a welding course. Evaluators given an array of facts, including the salient fact that the students who passed had a light course load, those who failed had a heavy course load.
 - Expt. Condition 1: 4 men pass, 4 women fail → evaluators identify gender as the reason the women failed.
 - Expt. Condition 2: 2 men and 2 women pass, 2 men and 2 women fail → evaluators identify course load as the reason for failure.

Letters of recommendation (and personal nominations) are enormously important for academics, in hiring, promotion, invitations to speak, fellowships, grants, and other honors and awards. Yet there are systematic differences in the letters of recommendation for women and for men (e.g., Trix & Penska 2003). This is not widely known among science and engineering faculties. Letters for women are shorter and contain fewer standout words (like “outstanding” or “ground-breaking” or “superstar”). They are more likely to mention women's personal lives, and in most cases, the mention of gender is explicit. Women are more likely to be compared to other women (a sure sign that this process is not gender blind). Letters for women express more doubt and contain more “grindstone” adjectives (“works hard,” “diligent,” etc.). In my own experience, women get asked to write tenure letters for women more often, and their letters are more likely to be discounted or ignored - unless, that is, they are negative, in which case they are given extra weight. That is, women are not reliable if they support other women (it is interpreted as solidarity), but if critical must be more discerning, since naturally they should be supporting other women. (In other words, women scientists are women first, scientists second.)

The presence of only a few women guarantees that bias will kick in. In studies of hiring practices, with artificial and matched resumes (Heilman 1980), it was found that women can succeed when they are more than 30% of the applicant pool, and that they are unlikely to succeed when less than 25%. This has obvious ramifications for job searches or tenure letters that include only one woman as a token on the short list.

As Virginia Valian describes in her book, “Why So Slow? The Advancement of Women,” expectations of men and women in our society are different, and those expectations - “gender schemas” - color our judgments, even those supposedly based on objective criteria. Schemas are expectations - often based on real characteristics - that help us interpret our surroundings. In this society, men are seen as capable of independent action, oriented to the task at hand, and acting on the basis of reason. Women are seen as nurturing, feeling, and prone to expressing feeling. Men *act*, women *feel and express feeling*. In the presence of schemas (e.g., in a profession dominated by men, like physics), applying gender schemas lead us to overrate men and underrate women.

Valian also describes how the “accumulation of disadvantage” - even small, seemingly minor disadvantages, can accumulate over a career to leave

women in a decidedly inferior position (conforming to the data). She illustrates this with a simulation (Martell, Lane & Emrich 1996) of a company with an 8-level hierarchy; even starting from 50/50 gender equity at the base level, a promotion system biased only 1% in favor of men quickly results in a top management tier that is 65% men.

This has been a very brief review of what is known from the sociology and psychology research, but enough, I hope, to show that this is not a mysterious problem. Rather, it is a well-understood and tractable problem.

Research on relevant issues:

Entitlement and self-image:

- *Major 1987*: women work harder and longer than men for the same pay, and will accept as fair a lower pay.
- *1991 Women's Tennis*: Seles suggests equal pay in tennis tournaments; Graf, Fernandez say publicly no, not necessary, we will seem greedy.
- *2007, Wimbledon, France* award equal prizes to women's and men's tennis champions (34 years after US Open)
- Women act for larger (community) good; men expect recompense.
- *Bowles, Babcock & Lai 2007*: less likely to hire woman who asks about money.
- *Sonnert and Holton 1996*: women rate themselves lower than men
- Compared to external ratings, men are likely to rate themselves above average, and women to rate themselves below average.
- Men wildly overestimate their future earnings.

Denial of Disadvantage

- *Clayton & Crosby 1992*: successful women do not work for the advancement of other women. Why? They want to believe in a meritocracy and that evaluations are objective; if not, it invalidates their success.

Remedies

Gender schemas resist change (and follow change). Changing them requires education, action and evaluation.

The first step toward change is to ***educate*** our colleagues about the impact of gender on evaluation and career progress. The National Academy of

Science's *Beyond Bias and Barriers* study summarizes the relevant research and interventions. Many NSF *Advance* projects² have online resources, and universities can develop effective methods to teach scientists the (social) scientific literature. Virginia Valian maintains a very useful annotated bibliography of relevant research (www.hunter.cuny.edu/genderequity/equityMaterials/Feb2008/annobib.pdf). You can assess your own comfort with gender equity (implicit.harvard.edu). *Advance* groups have also developed very effective advice concerning job searches³. It is essential to actually *search* for candidates rather than simply reviewing incoming resumes, and to be prepared to deal creatively with the dual career issue.

You can educate your colleagues about, for example, how to write letters of recommendation (Trix & Penska 2003). You can teach students about teaching evaluations, which are more negative for women faculty.

Change happens when leaders take *action*. Leaders need to establish norms. For example, make sure colloquia, meetings, prizes, job interviews, etc., involve the appropriate fraction of women. Find and promote talented women. Articulate the issues and provide training as needed. Hold people accountable. Strong leadership can make a big difference.

- *Brown & Geis 1984*: pre-validation of leaders; videotape of student evaluators believe their judgment was independent of introduction - but it was not. However, students responded positively to validation within the introductions.

Learn to be effective (from organizational development literature) in taking the message forward.

Information and mentoring are essential. A mentoring program at the Johns Hopkins Medical Institutions dramatically improved the tenure rate for women assistant professors (and also, by the way, for men who took part in the program - just one example of what's better for women is often better for men).

Other issues are more subtle. In many fields, the climate for women is inhospitable. Cultural values unrelated to ability or performance nonetheless

² *Advance* is an NSF program intended to transform academic institutions with respect to women in science. Nineteen institutions and consortia have been given *Advance* grants.

³ For example, www.washington.edu/admin/eoo/forms/ftk_01.html.

dominate perceptions of quality (e.g., arrogance, assertiveness, aggressiveness), and indeed may repel women from the profession. The University of Michigan *Advance* project has developed theatre performances that address this very effectively, and have been presented to national meetings of physicists, chemists, the National Science Foundation, Harvard University, and many others.

Evaluation will show which approaches are the most effective. Proven strategies should be adopted widely. Some of the *Advance* program toolkits (e.g., for hiring, retention, etc.) are well suited to use in a large number of institutions.

Summary

Data show the problem. Theory explains why it is pervasive.
Good intentions are not enough.

The main problem is our perception of women being less good than men, when objective (gender-blind) review says otherwise (e.g., orchestra auditions, resumes, etc.). Women are not automatically seen as leaders, or even as competent. Yet even this can be changed, by external validation by accepted authorities (usually men). For example, introducing a speaker with a well-thought out review of their status establishes that status in the audience's mind. Similarly, appointing suitable women to positions of leadership can have the effect of educating the community that they are deserving of those positions.

The key point is that change - toward greater equity and thus a higher level of excellence - takes positive intervention. It will not happen without action.