# 4 How Careers Progress for Different Groups: Observational Data and Alternate Accounts

In chapter 3 we presented laboratory data and results from field studies to support our two-part hypothesis about the evaluations of people who differ in terms of gender, race, and ethnicity. We concluded that gender, race, and ethnicity schemas result in a systematic small undervaluation of women and non-Whites and a systematic small overvaluation of White men in professional settings. The examples of undervaluation—such as being ignored, or rated lower, or given fewer opportunities to excel—are small but, we argue, frequent. Those small differences in advantage and disadvantage add up over time—like interest on investments—to create disparities in outcomes: molehills become mountains. Given our conclusions from the studies we reviewed in chapter 3, we expect to see differences in observational data as well.

In the first part of this chapter we examine hiring, salaries, and related outcomes. In the second part we consider additional hypotheses to those we presented in chapter 3. The underrepresentation of women and people of color in positions of power and prestige is a complex phenomenon that no single factor can explain. In some cases we think the additional hypotheses have merit and can be added to our interpretation in chapter 3 without supplanting it, while in other cases we think the hypotheses have at best weak evidence in their favor.

## Generalizations

Some generalizations hold across different samples.

*First*, academia is not unique among the professions. White women and men and women of color fare worse than White men in law, medicine, and business: they make less money and they advance more slowly. Because of the similarity across professions, and because very large samples are available for fields outside academia, we draw on the accumulated data.

*Second*, academia in the United States is not unique among the world's colleges and universities: women fare worse than men; minority groups fare worse than majority groups.

*Third*, the data are similar across different samples: differences in careers start out small and get larger as careers progress. The finding of growing disparities over individuals' careers is expected if some individuals can accumulate advantage more rapidly than others.

## Methodological Issues

In examining the reasons for the gender and race gaps that we observe, we face a central problem at the outset. It is impossible to include only variables that are not themselves a product of how people are treated. This is an important limitation, though it is often not considered in studies that "control" for the effect of variables that actually partially construct and reflect gender and/or race. Consider the following scenario. Whites are more successful in obtaining grants than are African Americans with the same qualifications (Ginther, Schaffer, Schnell, Masimore, Liu, Haak, & Kington, 2011); in addition, there is some evidence that female African Americans fare worse than male African Americans (Ginther, Kahn, & Schaffer, 2016). Now imagine that grant funding predicts salary (see, e.g., Kelly & Grant, 2012; Melguizo & Strober, 2007).

If we use grant funding as a predictor of salaries or as a factor we "control" for, we might diminish or eliminate the salary difference between African Americans and Whites. However, in using grant funding as a variable, we are using a measure that may itself be a product of inappropriate and disparate treatment. A finding that salary differences are reduced once funding is taken into account can make it seem as if a puzzle has been solved, when in fact the heart of the puzzle is still there, just further back in the causal chain. In other cases, we may be in the dark about whether a variable is the product of unequal treatment. For example, women are less likely to work in more remunerative areas, like engineering and economics, than men. Choice of field may be a preference but it may also or alternatively be a product of earlier disparate treatment.

The absence of a gap, by itself, does not entail the absence of unequal treatment, nor does the presence of an unexplained gap, by itself, entail the presence of unequal treatment.[1] For example, in the absence of a gap, women and men could fare equally well in a field because the women who remain are more highly selected than the men who remain in the field. Conversely, an unexplained disparity could be due to a factor that has not been measured.

Different studies consider different measures (such as prestige of degree-granting institution and propensity to compete) but seldom include estimates of the importance of each measure to the outcome (salary, rate of advancement, etc.), or the extent to which different groups (men and women, African Americans and Whites) vary on the measure, or the effect of the disparity on the outcome.

Finally, with a wealth of partial data—studies that are incommensurable because they do not measure the same variables—comes a wealth of competing interpretations of the data, depending on the theoretical framework of the investigator. We try to steer our way through the data by relying in part on the experiments we discussed in chapter 3.

## Data on Professional Trajectories

With the methodological cautions in mind, we first consider some representative large- and medium-scale findings in order to illustrate real-world gender and race and ethnicity differences. We then present other data about gender, race, and ethnicity in academia.

### Study 1—Wage Growth as Careers Progress

We begin with large-scale observational data from the United Kingdom on gender disparity (Manning & Swaffield, 2008). The study examined wage growth—how salaries increase over time—and shows (as do many studies) that the disparity between men's and women's wages develops gradually over people's careers. The United Kingdom data came primarily from the British Household Panel Study that covered the period from 1991 to 2002, and data from the New Earnings Survey, for a sample of more than 10,000 observations on roughly 3,500 people.

In this study, men and women had similar earnings when they first entered the labor market, but women's wages grew more slowly than men's, with the result that their incomes after ten years were lower than men's. Some of the gap could be explained by gender differences in time out from the labor market and other factors, but women were about 8% behind men even if they had been continuously employed full-time, had no children, did not plan to have any children, and had similar personality characteristics (Manning & Swaffield, 2008).

The British study's findings are similar to those in the United States: women progress more slowly than men. In some cases that have been analyzed, the slower progression can be accounted for completely by the variables that the investigators have used, such as in a study of MBAs (Bertrand, Goldin, & Katz, 2010), and in other cases there remains a gap which is not explained. (For a comprehensive review, see Blau & Kahn, 2017.)

### Study 2—Salary

Our second example comes from the United States. In 2016, the National Science Board (which is a federal group that establishes policies of the NSF), in a report titled *Science and Engineering Indicators 2016*, compared salaries of men and women, and different racial and ethnic groups, in science and engineering fields (where science includes psychology and social sciences as well as biological, physical, and computing sciences). The most recent available data were from 2013. Looking across all employment sectors, the Board reports that women make less money than men, and non-Asian minorities make less money than Whites and Asians, whether the most advanced degree was a BA, an MA, or a PhD. (Sample sizes are not reported for these comparisons, but other data tables suggest that roughly four million data points exist for the entire sample.)

For PhDs, without any controls, women made 21.4% less than men. Non-Asian minorities with PhDs averaged 16% less money than did Whites and Asians. After controlling for field of degree, occupational category, employment sector, years since PhD, years since degree squared, Carnegie classification of the school that awarded the PhD, and status (private or public) of the institution awarding the PhD, the gender disparity was reduced from 21.4% to 7.4%.[2] Although these data are not for academia alone, the analyses found disparities when controlling for sector of employment. The minority disparity, when the same controls were introduced, was reduced from 16.2% to 5.7%.[3] Whether all the controls introduced were appropriate is arguable. For example, the sector in which an individual is employed might itself be determined by employer preferences rather than employee preferences.

After adding further controls for citizenship, race and ethnic status (for gender), sex (for minorities), marital status, disability status, number of children living in the household, geographic region, and parental education, the gender disparity was slightly reduced, from 7.4% to 6.9%,[2] and the minority disparity was reduced from 5.7% to 4.8%.[3] Whether the third model's controls are appropriate is also arguable, but they reduced the disparities only by small amounts, and there continued to be disparities.[4]

Gender disparities in salary exist among both science and nonscience faculty, even after controls for productivity and teaching are included, although the patterns are somewhat different in the two broad areas (Kelly & Grant, 2012, using national data from 8,350 full-time faculty in 2003–2004). For an example of one female scientist's attempt to increase her salary, see box 4.1. Married and single women without children earn less

**Box 4.1**
**One Way Salaries Do Not Progress**

The year was 2005. The place was a basic science department in a major midwestern medical school. The scene was a meeting between Karen, an associate professor with tenure, and John, the chair of her department. Karen had two children; her husband was a stay-at-home father. Karen wanted an increase in her salary so that she would make the same amount of money that her colleague Phil, who recently got a raise, made. She thought that they were comparable. She was productive, she had grant support, she sat on an NIH study section. John said, "When your CV looks like Phil's, then we can talk about a salary increase." Karen thought their CVs *were* comparable, which was why she had approached the chair to begin with. John went on to say, "Or if you get an offer from somewhere else, come see me." Karen knew it was common for chairs and deans to adopt the approach of responding to outside offers, but she thought it was not good policy, since it encouraged people to look elsewhere, was distracting, and wasted time. Finally, John said, "Anyway, Phil's wife was making more money than he was, so I had to give him a raise." To that, Karen had no answer. In 2016, Karen decided to take early retirement, despite her love for her research and her success in obtaining grants to fund the research. Her department had changed leadership but the underlying problems remained.

than married fathers, but the disparity is greater in nonscience than science fields. Individual fields differ somewhat: business schools, for example, do not show a gender disparity once subfield and rank are accounted for (Sutanto, Bell, Fei, & Scott, 2014). Gender disparities in salary were not found across the board in all ranks and all fields (Ceci et al., 2014). However, medical schools show gender disparities in young faculty's salaries even after a broad range of controls is included (Jagsi, Griffith, Stewart, Sambuco, DeCastro, & Ubel, 2013).

To summarize, in 2013, women with PhDs in the sciences (broadly construed) made 7% less money than comparable men. Although gender disparities are common in a number of fields, they are not universal. Non-Asian minorities made 5% less.

### Study 3—Progress through the Ranks (Spain)

The third study examines promotion in academia in Spain (Anghel, de la Rica, & Dolado, 2011). These data are similar to U.S. data and show that the phenomenon is not limited to the United States. Near the beginnings of their careers men and women are much more similar than they are ten years after they have begun a career—even after one controls for a number of variables.

There are no gender differences in promotion to associate professor, whether one compares men and women in the aggregate or by specialization. However, there are gender differences in promotion from associate to full professor. When comparing men and women of the same age, with the same amount of time since their PhD, the same field of knowledge, the same recent academic publication record, and the same number of dissertations directed, men were 2.5 times more likely to be promoted than women.

There are also gender differences in the effect of having children on promotion from associate to full professor. When comparing men and women with the same personal and professional characteristics and the same productivity, fathers are 4 times more likely to be promoted to full professor than mothers. These data are reminiscent of those for the United States, showing the penalty that women with children face in universities (Mason & Goulden, 2004; Mason, Wolfinger, & Goulden, 2013), in corporations (Correll, Benard, & Paik, 2007), and in experimental simulations (Benard & Correll, 2010; Correll et al., 2007).

## Interim Conclusion

Our interim conclusion is that, as the experimental data suggest, women and racial-ethnic minorities fare worse than White men in compensation and advancement. What about other differences like sexual identity or disability status? We cannot examine observational data about these important issues because those characteristics are typically not recorded in the large data sets that are available, and they are usually not available in institutional data sources for reasons of privacy. For that reason it is especially helpful that some experimental studies do exist.

## Representation among Doctoral Degree Recipients

Who earns a doctoral degree? Although different tables from the NSF present slightly different numbers, the overall picture for different sexes and different races and ethnicities is similar. We concentrate on degrees to U.S. citizens and permanent residents, even though temporary residents currently obtain more than a quarter of all doctoral degrees and more than 50% of doctoral degrees in fields like engineering and economics. Women in all race and ethnic groups, African Americans, and Hispanics are represented more strongly among citizens and permanent residents than among temporary residents.

In 2015, about a quarter of doctorate recipients acquired their degrees in education, the humanities, or other nonscience fields. Most degrees are in the natural and social sciences, mathematics and computer sciences, and engineering.[5] Among U.S. citizens and permanent residents, as table 4.1 shows, Whites received most doctoral degrees, followed by Asians, Hispanics, and African Americans. The remainder consists of American Indians, native Hawaiians, those reporting more than one race, and those of unknown origin.[6] Among the different groups, Whites and Asians are overrepresented compared to their percentages of the population between ages 25 and 64, while African Americans and Hispanics are underrepresented.[7,8] Women and men are roughly equally represented among Whites, but men are underrepresented among African Americans, Asians,[5,9] and Hispanics.[6,7]

In all fields, from engineering to education, the percentage of women among doctorate recipients has slowly and steadily increased, partly because more women are receiving degrees and partly because fewer men are. For example, in 1995, women were 12% of PhD recipients in engineering. By 2015, that figure had risen to 23%. Women even became more numerous

**Table 4.1**
Percentage data for PhDs who are U.S. citizens or permanent residents in all fields (science and nonscience alike) in 2015, along with 2014 estimated census data for adults ages 25–64 in each ethnicity group; groups are mutually exclusive

| | Percentage of 2014 population[a] | Percentage of 2015 PhD recipients[b] | Percentage of PhD recipients who are male[b] |
|---|---|---|---|
| Whites | 63.3 | 72.3 | 51 |
| African Americans | 12.3 | 6.5 | 36 |
| Hispanics | 16.3 | 7.0 | 46 |
| Asians | 5.8 | 8.7 | 44[c] |
| All others | 2.2 | 5.5 | 51 |
| All groups | 99.9 | 100 | 49 |

[a]See note 7. [b]See note 6. [c]See note 9.

in some fields where they were already numerous. For example, in 1995, women were 61% of PhD degrees in education. By 2015, that figure was even higher at 68%.[10] Despite that increase, women's degrees are less likely than men's to come from prestigious departments (Weeden, Thébaud, & Gelbgiser, 2017).

In contrast, the percentage of underrepresented minorities barely moved over that 20-year period. For example, in 1995, African Americans received 2% of conferred doctorates in engineering; in 2015, they received 3.9%.[11] Similarly, African Americans received 3% of doctorates in the life sciences in 1995; that rose to only 5.3% in 2015.[10]

## Representation within Academic Institutions

After receiving their degrees, individuals who take jobs in academia achieve positions at institutions that vary in pay and prestige. The most prestigious institutions are those classified by the Carnegie Foundation as having very high research activity. In this section we investigate whether gender and ethnicity are related to having a job at the most prestigious institutions—those labeled as having *very high* research activity and those labeled as having *high* research activity. The most recent data available from the NSF are from 2013 and are limited to science, engineering, and health doctorates.[12] Table 4.2 displays the data. These data include all researchers, regardless of citizenship status. (Most PhDs are bestowed at institutions with very high research activity.)

**Table 4.2**
Proportional representation of men and women faculty, and different racial and ethnic groups, at institutions rated as having very high or high research activity

| | Total sample size | Proportion (of group) at *very high* research activity institutions | Proportion at *high* research activity institutions |
|---|---|---|---|
| Male | 163,600 | .47 | .16 |
| Female | 89,800 | .44 | .13 |
| Total | 253,400 | .46 | .15 |
| White | 187,800 | .46 | .15 |
| Asian | 41,500 | .53 | .16 |
| African American | 8,800 | .34 | .14 |
| Hispanic | 11,200 | .45 | .13 |
| All others | 4,600 | .39 | .11 |
| Total | 253,900 | .46 | .15 |
| With disability | 19,500 | .43 | .16 |
| Without disability | 234,000 | .47 | .15 |
| Total | 253,500 | .46 | .15 |

For gender, our calculations from NSF show significantly lower representation of women than men at *very high* (44% vs. 47%) and *high* (13% vs. 16%) research active institutions.[13] (Although the differences in percentages are small, the very large sample size provides enough power to determine that the difference is significant.) Women in science fare worse than men, in the sense that they are significantly less likely to be employed at the most prestigious academic institutions.

When we compare different ethnic groups,[14] the picture is more extreme. At very high research activity institutions, Asians are better represented than Whites, and Whites are marginally better represented than Hispanics and better represented than African Americans and all others. At high research activity institutions, Asians are again better represented than Whites and Whites are better represented than any other group. African Americans have a lower representation than any other group in institutions with very high and high research activity. In sum, African Americans fare particularly poorly in academic positions and Asians fare particularly well.

For individuals with and without disabilities, representation differs at the two types of institution. Individuals with disabilities have lower representation at institutions with very high research activity, and higher representation at institutions with high research activity.[15]

## Rank and Tenure

One marker of success is the speed by which one attains the top two faculty ranks: full professor (the highest rank except for named professorships and distinguished professorships) and associate professor. NSF provides data for science, engineering, and health fields. We compared only men and women who were less than ten years post-PhD in 2013.[16] We also only included individuals who are in the ranks of assistant, associate, or full professor in the denominator. Since a higher percentage of women than men are in instructor or lecturer positions, this is a conservative estimate of the gender disparity in rising through the ranks. The data are shown in table 4.3.[17]

Across all three fields (science, engineering, and health), men were significantly more likely to achieve the top two ranks within ten years postdegree than women were: 28% versus 24%.[18] For science fields as a whole, men were also significantly more likely to achieve full and associate professor status

**Table 4.3**

Proportions of men and women at full and associate professor ranks within ten years post-PhD

| | All fields | Science | Engineering | Health |
|---|---|---|---|---|
| Proportion of field that is male | .55 | .54 | .76 | .27 |
| Male | .28 | .27 | .28 | .33* |
| Female | .24* | .22* | .21* | .37 |
| Sample N | 57,600 | 44,000 | 7,600 | 5,900 |

*Note*. The proportion of a field that is male is calculated from all those with degrees less than ten years old who are employed in four-year educational institutions in any capacity. The denominator for proportions of men and women at senior ranks includes full, associate, and assistant ranks, but excludes "instructor or lecturer," "all other faculty," and "rank not applicable" individuals (N excluded = 43,900 for all fields), in order to restrict the comparison to those with ranked positions. Comparisons with an asterisk are statistically significant. For engineers, data for full professors were too sparse to be reported; only associate professors comprise the numerator. See also note 19.

than were women: 27% of men versus 22% of women.[18] For engineering, the comparable percentages (for associate professor level only, due to sparse data for full professors) are 27% versus 21%, also a significant difference.[18] For health, however, the percentages are reversed: 33% of men versus 37% of women achieve the top two ranks within ten years, again a significant difference.[18,19]

To investigate different science fields, we conducted individual analyses for the percentages of individuals who reached the rank of associate professor within ten years, as shown in table 4.4. (In most fields, NSF does not list the number of individuals at the full professor rank, because there are so few.) The fields are presented in descending order of the proportion of men in the entire field (including all ranks). In math, physical sciences, biological and agricultural sciences, and social sciences (excluding psychology), men are more successful than women.[20] In computer science and psychology, men and women are equally successful.[21]

To summarize our findings about success in reaching the top two ranks in academia within ten years in 2013, our data are limited to different fields within science. We find that in five fields (engineering, math and

**Table 4.4**
Proportions of men and women at associate professor rank within ten years post-PhD in individual science fields

| | Computer science | Math | Physical sciences | Biological/ agricultural sciences | Social sciences (excluding psychology) | Psychology |
|---|---|---|---|---|---|---|
| Proportion of field that is male | .77 | .69 | .67 | .53 | .51 | .32 |
| Male | .44 | .24 | .18 | .17 | .28 | .25 |
| Female | .40 | .08* | .16* | .10* | .25* | .25 |
| Sample N | 2,300 | 4,100 | 5,800 | 10,900 | 12,800 | 6,800 |

*Note*. Comparisons with an asterisk are statistically significant. Fields are ordered by the proportion of the field that is male and thus include all those with degrees less than ten years old who are employed in four-year educational institutions in any capacity. The N includes only associate and assistant professors; it excludes "instructor or lecturer," "all other faculty," and "rank not applicable" individuals, in order to restrict the comparison to those with ranked positions. It also excludes full professors because the N is so small that it often does not appear in NSF tables. See also note 21.

statistics, physical sciences, biological and agricultural sciences, and social sciences) men are more successful. In two fields (computer science and psychology) men and women do not differ. In one field (health) women are more successful than men. (Since there are two fields in which the sexes do not differ, and one in which women are more successful, differing patterns of time away from the job between men and women are unlikely to account for the differences we observe.)

In general, women fare best in achieving associate and full professorship status in the two fields with the highest percentage of women (health and psychology), but they fare as well as men in computer science. In three of the four fields that have been characterized as "math intensive" (engineering, physical science, and math), women are less successful than men, but that is also true for other fields that are not generally classified as math intensive (biological and agricultural sciences and social sciences).

By restricting our analysis to individuals who are less than ten years postdegree, we have reduced (but of course not eliminated) the possible extent of variation in qualifications. Disparities are considerably greater between 10 and 19 years postdegree, as one would expect from the accumulation of advantage and disadvantage, but as one would also expect because of the many variables (such as matching of outside offers) that could produce such disparities.

We performed the same gender comparisons for achievement of tenure within ten years for 2013 and present the data in table 4.5.[22] For this comparison we excluded individuals who are not on a tenure track or to whom tenure does not apply. We are thus comparing only men and women who have achieved tenure out of all men and women with tenure plus those without tenure but on the tenure track. Since women are overrepresented among those not on a tenure track and those in positions to which tenure does not apply, the gender disparities in achievement of tenure are conservative estimates.

Tenure and rank are correlated (those with tenure are likely to be in the top two ranks, and those in the top two ranks are likely to have tenure), so one would expect gender disparities in tenure. Although the differences across all fields are numerically extremely small, men are significantly more successful than women are (26% of men vs. 24% of women); the same holds for science as a whole (26% of men vs. 23% of women). In engineering, however, there is no difference in tenure rates, even though women lag behind men

**Table 4.5**
Proportions of men and women who have achieved tenure within ten years post-PhD for all fields, science, engineering, and health

| | All fields | Science | Engineering | Health |
|---|---|---|---|---|
| Proportion of field that is male | .55 | .54 | .76 | .27 |
| Male | .26 | .26 | .24 | .25* |
| Female | .24* | .23* | .25 | .33 |
| Sample N | 44,900 | 34,300 | 6,600 | 4,200 |

*Note*. Comparisons with an asterisk are statistically significant. The proportion of the field that is male includes all those with degrees less than ten years old who are employed in four-year educational institutions in any capacity. For tenure, the N excludes "not on tenure track" and "tenure not applicable" individuals (N excluded = 56,600 for all fields, N excluded = 46,300 for science fields).

in becoming associate and full professors (table 4.3). (The reasons for the differences in rank and tenure success are not clear.) In health, women are significantly advantaged in tenure, as they were for rank (table 4.3).[23]

We also examine broad fields within science; the data are shown in table 4.6.[24] In math and statistics, the physical sciences, and the biological sciences, men are significantly more likely to achieve tenure within ten years postdegree than are women, mirroring the results for achievement of associate professor status. These fields include ones where the total percentage of men in the field is high (e.g., mathematics and statistics) and one where it is similar to the percentage of women (biological and agricultural sciences). In computer science and the social sciences, there are no significant gender differences. In psychology, where women outnumber men, women are significantly advantaged; the same holds for health. In engineering, where men outnumber women 3:1, the two groups are nevertheless equal in tenure rates, similar to the social sciences, where men only slightly outnumber women (51%).

The data in tables 4.5 and 4.6 do not show a tight connection between the percentage of women in a field and their achievement of tenure within ten years. In both engineering and computer science, where women are not quite 25% of the field, there are nevertheless no differences in tenure rates. In biology, where young women are almost as numerous as young

**Table 4.6**
Proportions of men and women who have achieved tenure within ten years for each broad field within science: biological/agricultural/life sciences; computer/information sciences; mathematics/statistics, physical sciences, psychology, social sciences (excluding psychology), engineering, and health

| | Computer science | Math | Physical sciences | Biological/ agricultural sciences | Social sciences (excluding psychology) | Psychology |
|---|---|---|---|---|---|---|
| Proportion of field that is male | .77 | .69 | .67 | .53 | .51 | .32 |
| Male | .44 | .26 | .21 | .20 | .28 | .24* |
| Female | .40 | .11* | .13* | .13* | .29 | .27 |
| Sample N | 2,300 | 3,200 | 4,900 | 7,600 | 11,200 | 5,000 |

*Note*. Fields are ordered by the proportion of the field that is male and include all those with degrees less than ten years old who are employed in four-year educational institutions in any capacity. Comparisons with an asterisk are statistically significant.

men, men are more successful. But in psychology, health, and social science, where women are more numerous, they are either more successful than men or equally successful.

Our data are generally, but not always, consistent for achievement of rank and tenure across fields in science, engineering, and health. In math, physical sciences, and biological and agricultural sciences, men fare better than women in both rank and tenure. In engineering and social sciences, men have an advantage in rank but are equal with women in tenure rates. In computer science, men and women are on a par in rank and tenure. In psychology, men and women are on a par in rank, but women are advantaged in tenure. Fields may differ in whether rank or tenure is the more difficult achievement.

Overall, we would sum up our findings for men and women in the sciences as showing that women are only more advantaged relative to men if they are extremely highly represented in a field (health and, for tenure, psychology). Otherwise, men are either more advantaged in achievement of rank and tenure, or on a par with women. We also note that if we included all individuals, such as instructors and lecturers, gender disparities would be greater.

Not all studies are consistent with our findings. The differences across studies are difficult to evaluate but may be due to differences in methods, data set, and types of analyses. One gender analysis separated individuals by broad field (life and social sciences vs. natural sciences, engineering, and math) and found that women fare worse than men in the life and social sciences but do as well as men in the natural sciences, after a number of characteristics, including marital status, are controlled for (see Ceci et al., 2014, for a summary). (We would argue that controlling for marital status is overcontrolling. If employers treat married men and women differently, that is a problem that should not be masked by using marriage as a control.) A study limited to junior faculty in the natural sciences and math at 14 universities found no gender differences in time to tenure or in retention, except in math (Kaminski & Geisler, 2012); no factors were controlled for. A study of faculty in the social sciences found differences in time to tenure (Box-Steffensmeier, Cunha, Varbanov, Hoh, Knisley, & Holmes, 2015).

Earlier we reviewed data on rank progression in universities in Spain showing slower progression for women. The phenomenon of slower rank progression for women also occurs in the Netherlands, as shown in a study of young researchers in three social science fields (psychology, economics, and behavioral and educational research) roughly ten years post-PhD (van den Besselaar & Sandström, 2016). Researcher productivity predicted rank, and men published more than women (but had equal citation rates), but gender had an effect above and beyond productivity and other controls (van den Besselaar & Sandström, 2016). At the beginning of their careers, the young researchers did not differ in productivity. Men's productivity grew faster than women's over the ten-year period. That finding is susceptible to varying interpretations that are not mutually exclusive: women may be given fewer resources, including mentoring and collaboration, than men; women may work less intensively than men; women may produce work of lower quality; or women may perceive that they must reach a higher standard.

## Summary

To conclude, women in the sciences are significantly less likely than men to have positions in universities with very high or high research activity; in several fields, they are significantly slower to achieve tenure within the first ten years post-PhD and are less likely than men to become an associate or full professor; they also make less money. African Americans are significantly less likely than Asians, Whites, or Hispanics to have positions in universities with very high research activity, while Hispanics are on a par with Whites. African Americans and other non-Asian minorities make less money than Whites and Asians.

In the workplace as a whole, including academia, men in general fare better than women in terms of promotion and salary, even after considering a number of factors that reduce gender disparities (Blau & Kahn, 2017). Similarly, White men fare better than non-Asian men and women of color.

Although we conclude that White men have more successful careers, as conventionally defined, than do White women or non-Asian men and women of color, we also reiterate that careers for White women in particular have markedly improved since 1970. That noticeable improvement can draw attention away from the stubborn inequalities that remain. We attribute the recalcitrance of those inequalities to two factors. One is the differences in evaluations of people, differences due to schemas that people hold about the abilities and performances of different groups, as described in chapter 3. The other factor is the absence of policies and procedures that will reduce the likelihood of incorrect evaluations, for which we recommend policy changes in later chapters. (We consider other possibilities in the next section.)

We also suggest that visible exceptions—very successful White women, very successful male and female African Americans—can disguise the inequalities that remain in the advancement of underrepresented groups. Our focus on the exceptions can mislead us into thinking that there is no problem; that logic implies that if there were a problem, then there would be no successful individuals from those groups. However, it is important to realize that the exceptions are exceptions, not the norm.

## Accounts of the Observational Data

As detailed in chapter 3, people evaluate White men differently from White women and people of color, even when performances are the same. Many of those evaluations are small daily occurrences that can be barely visible but that mount up over time to accumulate advantage (White men) or disadvantage (White women and non-Asian men and women of color). Although the observational and experimental studies that we have cited support our explanation, there are other accounts that have been provided for the observational data.

We consider in turn seven additional accounts of the patterns of career advancement that we have detailed.[25] As we noted earlier, we think some of the accounts have merit and can be interpreted within the framework we developed in chapter 3. For others, the accounts adduce facts that are not well established.

### Seven Accounts

1. Demographic inertia, based on the fact that there are so many more older White men than women, or so many more older Whites and Asians than people of color in various samples or populations:

• The problem will take care of itself in time; there is nothing that academia needs to do differently.

2. Disparities in distribution of *extra*-academic responsibilities, such as childcare within heterosexual couples or community work on the part of White women and African American and Hispanic men and women:

• There are societal and cultural issues preventing equal advancement that are outside the purview of academia; there is nothing that academia needs to do differently, but it could provide benefits making childcare more available and accessible.

3. Disparities in ability:

• Members of different social groups have different abilities; as a result, members of different social groups go into different fields and have different levels of success; there is nothing that academia needs to do differently.

4. Disparities in productivity:

• Members of social groups differ in their average productivity; as a result, members of different social groups have different levels of success; there is nothing that academia needs to do differently, but it could "mentor" groups to increase productivity.

5. Disparities in interest:

• Members of different social groups have different interests; as a result, members of different social groups go into different fields and have different levels of success; there is nothing that academia needs to do differently.

6. Disparities in attempts to negotiate, compete, or self-promote:

• Members of different social groups have different skills or interest in negotiating, competing, or self-promoting; as a result, members of different social groups go into different fields and have different levels of success; there is nothing that academia needs to do differently, but it could "mentor" individuals from those groups to increase their skills.

7. Overt bias:

• Academia needs to reduce bias.

## 1. Demographic Inertia

In the case of gender, the account draws on the fact that, in the past, the "pipeline"[26]—especially in the natural sciences and engineering—had many fewer women compared to men. Because White men in the past achieved advanced degrees at rates much higher than did White women or men and women of color, they automatically dominate the higher ranks of academia. According to this hypothesis, men are so numerous in higher ranks because there are a lot of older men around. Women are concentrated in lower ranks because there are so many young women (compared to older women) around. Under this hypothesis, the small number of women full professors is the consequence of a relatively recent increase of young women in the sciences (Hargens & Long, 2002). The same hypothesis can be advanced to explain the low representation of African Americans and Hispanics, whose representation among PhD recipients is still very low, and considerably below their representation in the population as a whole.

There is a statistical kernel of truth here. If an area continues to award only 20% of its PhDs to women, that field will never have more than 20% women at the highest ranks unless men drop out in very large numbers.[27] Similarly, if Hispanics and Blacks remain only 6% of PhD recipients in the sciences, it will be a very long time indeed before their representation among full professors increases.

Despite the kernel of truth, the inertia hypothesis does not account for all of the facts. With respect to gender, women's movement through transitions does not match what demographic inertia would predict (Ceci, Ginther, Kahn, & Williams, 2014, who also suggest that subfield differences exist; Shaw & Stanton, 2012; Thomas, Poole, & Herbers, 2015). Inertia predicts a higher percentage of women in graduate school, postdoc positions, and assistant professorships than is in fact the case,[28] though there may no longer be a drop-off between a bachelor's degree and a doctorate (Miller & Wai, 2015). Even in fields where women are well-represented, such as psychology, women do not succeed at the same rate as men. For example, women leave psychology after graduate school at roughly the same rate that they leave physics (Shaw & Stanton, 2012). Having a large number of women with the appropriate degrees does not guarantee that they will remain in that field. The "pipeline" leaks women. Only if policies and procedures change at every point in the pipeline, from recruitment to retention to promotion, will the leakage stop (Thomas et al., 2015).

We conclude that demographic inertia alone cannot account for the fact that different demographic groups progress through their careers at different rates.

### 2. Extra-academic Responsibilities

In most dual-earner two-sex households, women and men do not equally divide their household and childcare responsibilities. Further, few universities provide adequate parental leave or childcare facilities for parents. As a result, according to this account, women do not have the necessary time to devote to becoming successful at work. This account predicts that women will leave the professions in greater numbers than men, and that those who remain in the professions will achieve less than men. We find a great deal of evidence to support the first prediction, but not the second.

Before we continue, we note an underlying assumption: achievement is importantly related to hours of work. It is obviously necessary to spend some amount of time working, but we do not know how much time is necessary or what the function is that describes the relation between time and achievement. It is even difficult to decide what counts as work. Does supervising graduate students count? Does reading other people's work count? Does going to a talk on a tangential topic count? Does data analysis count? Does thinking count? Or does only writing count? If they all count as work, is more always better? There is no literature on the relation between time spent working (however defined) and achievement, so the assumption underlying the hypothesis is speculation. We note, however, that creativity benefits from spending some time away from the explicit effort of working on a problem—the "incubation" effect (Sio & Ormerod, 2009).

One kernel of truth to this hypothesis is that women who live with men do more than their share of housework and childcare (Bianchi, Sayer, Milkie, & Robinson, 2012). Among all heterosexual married couples with children in 2009–2010, women averaged 18.3 hours of housework and 13.7 hours of childcare per week. The comparable figures for men were 9.5 and 7.2 (Bianchi et al., 2012). Thus, women were performing almost twice as much housework and childcare as men. Lack of parity holds no matter how much each partner earns. Another kernel of truth is that not enough universities have adequate childcare policies, and those that have them do not necessarily implement them (Ahmad, 2016). Parents are forced to find their own solutions.

The first prediction of this hypothesis is borne out. Women leave the full-time workforce, either to no paid employment or to part-time paid employment, at rates higher than men (Ceci et al., 2014; Goulden, Mason, & Frasch, 2011; for a review, see Ahmad, 2016). Women's exit from the workforce creates losses for them (to the extent that their work ambitions cannot be fulfilled), the institutions where they had worked, the institutions that trained them, and the fields in which they had worked.

The second part of the hypothesis is not borne out. Women who remain full-time in academia work the same number of hours as men, though they do not necessarily divide their time at work in the same way (Ceci et al., 2014). Women appear to spend the same amount of time in research as men around the world (Bentley & Kyvik, 2013). There are few if any productivity differences between women with and without children, though the findings are not completely consistent across studies (Ceci et al., 2014; Fox, 2005; Schroen, Brownstein, & Sheldon, 2004). To the extent that women publish less than men, it does not appear to be due to the presence or absence of children (though men with children may publish more than men without children; Ceci et al., 2014). Women without children, according to this account, should advance as quickly as men, but that is also not the case (Kelly & Grant, 2012, though see Kahn & Ginther, 2017, who outline data suggesting that never-married women without children earn more than never-married men without children, once a suite of controls is introduced; for married men and women with children, controls reduce but do not eliminate a wage gap in the other direction).

Heterosexual women's ability to continue publishing their work despite their greater responsibilities at home may be due to their being a highly selected group. Almost by definition, people who survive in an environment that seems stressful are different from people who do not.

A different kind of responsibility is engagement with one's community. This is work that White women and men and women of color take on more than do White men (see O'Meara, Sandmann, Saltmarsh, & Giles, 2011, for review and discussion). Engagement with one's community can also be reflected in one's work, by, for example, codesigning studies with community members or enlisting community members as participants. Whether community work hinders faculty's research and teaching careers in some ways is not known, but institutions should value efforts by faculty to benefit their communities.

If institutions improved both the availability and accessibility of parental leave and childcare policies (and ensured that parents who took leave cared for their children), there would be less loss of talent and parents would benefit. For the women who do remain in academia, the data do not suggest that having children affects productivity, though having children may affect how women are evaluated and thus how much they are paid.

### 3. Ability

An account that attributes different professional outcomes to different abilities is periodically popular. In the area of gender, mean differences between males and females on most cognitive tasks are small or nonexistent. Attention has thus shifted from average differences to the ratio of males to females among those most talented at math. The argument here is that men are disproportionately found at the upper tail of scores on standardized math tests and are therefore more likely to excel in fields that require math.

Young males *are* disproportionately represented at the high end of standardized tests of math compared with young females, but the ratio of males to females in the top tenth of a percent of the math SAT shrank from 13.5 in a cohort tested in 1981–1985 to 2.53 in a cohort tested in 2011–2015 (Makel, Wai, Peairs, & Putallaz, 2016).[29] Thus, there is nothing immutable about the distribution of males and females with respect to math skills. Among adults, men are much more highly represented than women in math departments and among winners of math prizes. Some data suggest that girls' math ability is not cultivated to the same extent as boys' (Ellison & Swanson, 2010).

Partially supporting the hypothesis that early evidence of strong math skills is related to later achievement are data showing that people at the uppermost tail of the math SAT distribution as 13-year-olds—the top 1%—have by ages 48–53 amassed a striking number of accomplishments in science and other fields, dramatically more than base rates would predict (Lubinski, Benbow, & Kell, 2014).

One cannot, however, infer from high accomplishments that being in the top 1% of a standardized math test at an early age is necessary, sufficient, or jointly necessary and sufficient for high achievement in the sciences. The lack of sufficiency is apparent from the fact that women superscorers were much more likely to be full-time homemakers than were men, women superscorers made less money than men, women superscorers had fewer achievements than men, and women superscorers were less likely to have careers in the natural sciences than men (Lubinski et al., 2014). It is not enough to be a high scorer.

It is also doubtful that super high math scores on the SAT are necessary for success in math and science fields. The average quantitative GRE score for PhD students in math-intensive fields is around the 75th percentile, suggesting that the much greater percentage of men than women in those fields cannot be attributed solely to differences in math facility (Ceci et al., 2014). (Whether exceptionally high achievement is linked to exceptionally high scores on such standardized tests is not known.) Further, personality variables (such as conscientiousness) predict students' grade point averages above and beyond their test scores (Noftle & Robins, 2007).

Finally, the appeal to differences in math ability as an explanation of the slower advancement of women in math and science fields cannot, presumably, be used to explain the slower advancement of women in the humanities, where women also do not fare as well as men (Ginther & Kahn, 2006).

At present, there are no race or ethnicity studies for academia parallel to these gender studies. As we noted earlier, African Americans are less successful than Whites in obtaining grant funding even after controlling for a variety of characteristics. Thus, while there may be average differences in cognitive test scores,[30] African Americans whose characteristics are equivalent to Whites' do not reap the same benefits.

We have two general criticisms of accounts based on ability differences. The first is that they oversimplify what is necessary for success. No single characteristic, whether it be mathematical or verbal skill,[31] creativity, persistence, hard work, or a conscientious personality, is necessary or sufficient for success in any field. Success is due to a wide range of qualities, including luck, and in any given case it is hard to know how to apportion credit. The second criticism is that ability accounts fail to consider the differential treatment that individuals from different demographic groups receive and the mutability of high scores. If we want people to succeed, we have to have high expectations for them and supply constructive criticism.

## 4. Productivity

Decisions about hiring, tenure, and promotion are nominally made on the basis of "quality," and quality is generally measured by two values, both of which are imperfectly connected to quality: quantity and citations. We discuss the extent of sex differences in both measures and the extent to which those sex differences in both measures are related to job success.

The most common measure of productivity is number of publications, in part because it is easy to tabulate. The measure has known limitations. For example, researchers can inflate the number of articles they publish by writing very short articles. Several short articles may not be as substantive as one longer article that ties together several approaches, but on a quantity metric they will count more. Work that is new or goes against current thinking may be difficult to publish and require more evidence to publish, especially if one works at an institution that is not among the most prestigious.

Another common measure of productivity is number of citations. This is a measure of the impact—for good or for ill—that one's work has had.[32] The impact a publication has is determined in part by the perceived quality of the work, the prestige of the journal or academic press where the publication appears, the prestige of the person's institution, the number of people in a field, and current fashion.

Someone who writes a manual that changes how medical examinations are performed has clearly produced something of value. That manual may have required a great deal of time, effort, and analysis. However, it may not count as a publication since it may neither be in a journal nor exist as a book. The established metrics miss that. As another example, someone might write a single article on Shakespeare that sparks interest on the part of directors, who now produce plays differently than they would have before the article was published. It is just a single article, and there may be few citations because the impact that the work has is on the performance of Shakespeare's plays rather than on Shakespearean scholarship. Again, the established metrics will fail to capture the value of the work.

**Gender Differences in Productivity** If we stick with the established metrics, despite their limitations, we can compare men and women on number of publications. There is enormous variability and inconsistency. In some fields, in some countries, at some ranks, at some time periods, women have produced fewer publications than men. In others, there have either been no gender differences or women have published more. One international study of millions of papers reports that women publish less than men and are cited less than men (Larivière, Ni, Gingras, Cronin, & Sugimoto, 2013).[33] What is almost always found is that a higher percentage of men than women are extra prolific.

The finding that women publish somewhat less than men often occurs in studies where type of institution is not controlled in the analysis. Women are disproportionately found in institutions with heavy teaching responsibilities, while men are disproportionately found in research-intensive institutions. After controlling for such factors, Xie and Shauman (2003) conclude that the gender disparity in productivity in science is an artifact of the gender differences in institutional home.

Other factors that have been found to play a role—albeit inconsistently—in gender disparities in number of publications include time spent in research versus teaching, subfield (Ceci et al., 2014), degree of specialization (Leahey, 2006), and the extent to which a field is resource intensive (Duch, Zeng, Sales-Pardo, Radicchi, Otis, Woodruff, & Amaral, 2012). The inconsistencies are more striking than the consistencies. Sometimes, for example, there are no differences in publication rates between male and female psychologists (Joy, 2006), and sometimes there are (Ceci et al., 2014).

Citations are, overall, less likely than publication counts to show gender disparities, though, again, there is inconsistency across studies, often because of inconsistency in what controls are used. Men tend to cite themselves more than women do, but even that finding is not universal across fields. One innovative suggestion for calculating impact via citations is to subtract years when scholars had no publications and to exclude self-citations; once that is accomplished, there are no gender differences in citations in ecology (Cameron, White, & Gray, 2016). Women are more likely than men to experience years with no publications (as are people who are working on a long-term project); women, as mentioned above, are in general less likely to cite themselves than are men.

**Importance of Gender Differences in Productivity to Job Success** The main question, however, is whether gender differences in productivity determine promotion and salary. This is an even harder question to answer than the question about sex disparities. Independent of gender, the data suggest that productivity is a factor, but only one among many others (Byrnes, 2007; Ceci et al., 2014). An analysis of assistant professors in psychology departments at highly ranked universities found that 25% of those tenured had published less than one paper per year (Byrnes, 2007). One Dutch study found that gender predicted rank above and beyond productivity (van den Besselaar & Sandström, 2016). Departments have many reasons for retaining and promoting faculty. Productivity is one, but not the only one. We do not know of studies comparing Whites and underrepresented minorities.

Our discussion of productivity thus far has concentrated on research productivity, although institutions of higher learning have a mission to educate students. Even in teaching-intensive schools, research productivity is increasingly the principal criterion for hiring, promotion, and tenure. Junior faculty are often informally told that being a bad teacher will hurt their chances for promotion or tenure, but being an outstanding teacher will not suffice to gain promotion or tenure. Research productivity is the primary measure. Since there is evidence that students rate women worse than they rate men as instructors, even after controlling for various factors that could be relevant (Boring, 2017, using data from France; Boring, Ottoboni, & Stark, 2016, examining data from France and the United States; MacNell, Driscoll, & Hunt, 2015, with U.S. data subsequently used by Boring et al.; Wagner, Rieger, & Voorvelt, 2016, with data from the Netherlands), women are at a disadvantage in teaching when schools take students' evaluations into account.

Other forms of productivity can be measured. In acute care hospitals, for example, one can measure the extent to which physicians follow best practices and protocols to reduce errors as well as patient mortality and readmission rates (Tsugawa, Jena, Figueroa, Orav, Blumenthal, & Jha, 2017). Although the differences in performance are small, they significantly favor women. Women follow best practices more than men, and fewer of their patients die or are readmitted with complications. Women are nonetheless paid less and take longer to achieve full professor status (for a summary, see Parks & Redberg, 2017).

## 5. Interest

As appeal to intellectual disparities has waned, appeal to interest disparities has waxed. According to the different-interests hypothesis, White women, and men and women of color, are less interested in success, traditionally defined, than White men are (e.g., Ceci & Williams, 2010; Lubinski et al., 2014). For that reason they do not pursue challenging and demanding occupations to the same degree, or, even if they are in such occupations, they do not pursue success. Differential interests have also been recruited to explain why women are underrepresented in math, computer science, chemistry, physics, and engineering (Kahn & Ginther, 2017).

We reject an explanation that relies on differential interests if those interests are portrayed as inherent and immutable. First, people's interests and choices can and do change. Second, and more importantly, interests, choices, and goals are affected by the expectations we have of others, the opportunities we give them to succeed and belong, and the information we provide about what different fields have to offer, as research we considered in chapter 3 suggests. Everyone's potential interests are malleable, and most fields are compatible with a range of interests.

Women are often characterized as interested in people while men are characterized as interested in things. This picture is wrong for several reasons: it is too limiting for both groups; it incorrectly places people and things as opposed, instead of complementary; and it oversimplifies the data giving rise to the characterization (Valian, 2014).

Women *are* more interested in helping people than are men, and women have more progressive political stances. Women see their goals, which include helping others, as incompatible with the goals of science (Diekman, Brown, Johnston, & Clark, 2010). Students perceive most natural sciences, math, and engineering as less compatible with goals involving intimacy and altruism than with goals involving power and achievement.

If students have goals that are communal, they are likely to see science as a bad fit. And women are more likely to have such goals than are men. Changing how work in the sciences is portrayed, to make its social benefits clear, is one way to ensure that people with a range of goals will enter it. Similarly, high achievement in any field might look more attractive if it clearly included ways of using the position to make changes one deems beneficial.

We note that congressmen (Iacus, King, & Porro, 2011, confirming Washington, 2008) and male judges (Glynn & Sen, 2015) are more progressive on women's issues if they have daughters than if they have sons. For judges, the effect is particularly strong for judges who have only one child, a daughter. It should not come as news that people change their interests, and that individuals' interests are affected by the people around them. But the idea that free choice (constrained by childbearing and child-rearing) determines people's professional lives has become popular (e.g., Schmidt, 2011; Williams & Ceci, 2015).

Interest in fields like computer science is affected by the environment (Cheryan, Plaut, Davies, & Steele, 2009); women's interest in the sciences is affected by whether they think they belong and can be a valued member of a field (e.g., Cheryan, Ziegler, & Jiang, 2017; Cheryan, Master, & Meltzoff, 2015; Cheryan & Plaut, 2010; Dasgupta & Stout, 2014; Smith, Lewis, Hawthorne, & Hodges, 2013); both men and women sabotage peers who excel in areas that are not typical for their sex (Rudman & Fairchild, 2004). All fields that show a sex disparity, including childcare, have this problem: people who identify with the sex that is less well represented feel they do not have as much right to be active in that area. It is hard for fathers to play a time-intensive, nurturing role with their children when so few fathers do.

Marie Wilson, speaking of women in politics, said, "You can't be what you can't see." Something like that happens in every field, for both men and women. Seeing people like oneself in a position opens up that position as a possibility. For people of color the situation is more complex. Even if they were to be represented among leaders at rates matching their representation in the population as a whole, no single underrepresented minority would be very numerous. It is difficult for people from small demographic groups to see themselves in many positions. However, it may not be necessary to see people exactly like oneself. A diverse mix of people in a position may be enough of a signal that the position is achievable. You can be what you can't see if you see enough diversity to think there is room for you, too.

One example of the impact of even minor discouragement was reported by the *The Boston Globe* (Bombardieri, 2005). A male economist told a group of students that professors needed to be aware of the great influence that positive or negative signals can have on their students. He said that as a student he had had a genuine interest in economics, but he also went into economics because he had been dissuaded from some other fields "by experiences where I lagged slightly and where I was made to feel inadequate." He described an occasion when he gave the wrong answer to a physics question and "the person who saw my answer looked on with a certain stunned belief that I could be so stupid." This is a fairly mild example of discouragement, in which there is no explicit negative content, yet it was sufficient to deter Lawrence Summers (a former president of Harvard University who speculated about reasons for women's lower representation in the sciences) from pursuing physics.

Maybe Lawrence Summers would have been a mediocre physicist. Or maybe he would have been a good one. We will never know. We will similarly never know how many women and underrepresented minorities will show interest in the natural sciences and math unless we provide them with the right "opportunity structures" and the perception that they belong. As one study concludes,

> Women's interests are fundamentally shaped by the culture of [STEM] fields. Just because women are excited to go into other fields does not mean that they would not have been equally excited to go into computer science, engineering, and physics if the cultures signaled to them that they belong there. (Cheryan et al., 2017, p. 22)

Similarly, interests are shaped by challenges. Men and women receive different responses at work (King, Botsford, Hebl, Kazama, Dawson, & Perkins, 2012). Male and female managers in the energy and health care industries, for example, reported equal numbers of opportunities to expand their capabilities, but men received *challenging* opportunities at higher rates than women did. Experimental work suggested that men and women articulated equal interest in challenging opportunities but that male managers acted to protect women from possible failure, thus denying them the challenges that might make them more competitive for higher level jobs.

In short, interests change and choices change. By changing people's environments, we can provide them with more opportunities.

## 6. Negotiation and Competition

We begin this section with two stories. The first story comes from Karen, in box 4.1. Karen had tried to negotiate but failed.

The second story comes from Matthew, an African American faculty member at a public research university, who was offered a distinguished chair at a well-known historically Black college. While the change would bring some losses (in the status of the institution employing him and in having fewer graduate students), it offered some important gratifications, in particular the chance to teach many gifted African American students. He talked with his chair about the offer and the possibility of a competing offer from his current institution. His chair said, "We only make counter-offers when people have offers from peer institutions. We don't think you'll accept this offer." Matthew accepted the other offer. He too had tried to negotiate but failed.

The hypothesis about negotiation is that White women lose out because they do not negotiate well and avoid competition. This hypothesis has only been proffered for gender disparities, not for race and ethnicity disparities. There is a basis for the hypothesis, in that women negotiate less than men (e.g., Babcock & Laschever, 2009) and engage in competition less in some but not all settings (Andersen, Ertac, Gneezy, List, & Maximiano, 2013; Croson & Gneezy, 2009; Mitchell & Hesli, 2013). Two questions to ask are why that is so and whether that accounts for women's slow advancement relative to men's. We suggest that women and men receive different responses to their attempts to negotiate and compete, and that those different responses affect their behavior. In discussing schemas, we reviewed experiments that show that women who are assertive pay a penalty in being disliked, and that liking is related to benefits (Heilman, 2012; Heilman & Okimoto, 2007; Ridgeway, 2011; Rudman & Phelan, 2008).

What holds for assertiveness also holds for negotiation: people do not respond as favorably to women as to men when they negotiate for themselves (Amanatullah & Tinsley, 2013; Bowles, Babcock, & Lai, 2007; Kulik & Olekalns, 2012). Thus, women don't negotiate in part because, when they try to, they are seen as self-aggrandizing or unlikable (Kray & Thompson, 2004; Ridgeway, 1982), and do not get what they want unless they have already achieved high status (Amanatullah & Tinsley, 2013). Women are more likely to negotiate for higher salary if the job description suggests that negotiation is possible, whereas men negotiate regardless of whether the job description suggests that negotiation is possible or not (Leibbrandt & List, 2014).

Women who negotiate on someone else's behalf are generally more successful than women who negotiate for themselves because in that case they are not promoting themselves; experience with negotiation helps, as does knowing ahead of time what the relevant range of possibilities is (Mazei, Hüffmeier, Freund, Stuhlmacher, Bilke, & Hertel, 2015). Since women in academia can only negotiate for their own salaries, are sometimes doing so at the beginning of their careers when they have had less experience, and generally have only a vague idea of what the relevant salary range is, they will be at a disadvantage compared to men.

There is almost no research on African Americans' or other minority groups' attempts to negotiate or success at negotiation. Audit studies suggest that African Americans are less successful in negotiating with car dealers (Ayres & Siegelman, 1995). Whites bid less online for baseball cards when the hand holding the card is black rather than white (Ayres, Banaji, & Jolls, 2015), and bid less often and with lower amounts for a portable playing device if the hand holding the device is black rather than white (Doleac & Stein, 2013).

Status affects negotiation. If someone has low status because of his or her gender, race, ethnicity, or rank, he or she cannot negotiate successfully in the same way that someone who has high status negotiates. High status legitimizes negotiation. Women are asked more often than men to occupy service roles (O'Meara, Kuvaeva, Nyunt, Waugaman, & Jackson, in press) and accept (O'Meara, Kuvaeva, & Nyunt, 2017), perhaps in part because they suspect they lack the status to refuse. A similar phenomenon holds for competition, as demonstrated by gender differences in competition in matrilineal versus patrilineal societies (Gneezy, Leonard, & List, 2009). In matrilineal societies women are considerably more competitive than they are in patrilineal societies. Further, in both kinds of societies, girls and boys are equally competitive until around age 13, at which point girls in patrilineal societies (but not in matrilineal societies) stop being as competitive as boys. As is the case with interests, negotiation and competition are influenced by the surrounding context.

Willingness to engage in negotiation and competition can be determined by many considerations, including expectations of winning and the perceived (and possibly actual) social acceptability of competing. Women's math performance suffers when they actively compete against men (Niederle & Vesterlund, 2010; for review of gender effects from an economics perspective, see Niederle, 2016), possibly due to stereotype threat (Ben-Zeev, Fein, & Inzlicht, 2005). In line with that suggestion, women are more likely to compete in an experimental setting if the task is a verbal task—where stereotype threat is absent—than if it is a math task; they also produce work with fewer errors than men if there is no time pressure (Shurchkov, 2012). The ratio of male to female job applicants varies significantly depending on fairly subtle differences in the ways jobs are framed and on differences among applicants that are not related to gender. For example, women are more likely to apply for a job in which their pay would be determined by their performance relative to others if they were to work as part of a team. Gender differences are more extreme among younger applicants than older applicants; among older applicants, men and women equally prefer a fixed wage. Women are also less likely than men to apply for jobs that seem "masculine" (Flory, Leibbrandt, & List, 2015).

In sum, men's and women's willingness to negotiate or compete is not all-or-nothing. If the environment sends signals that they can benefit from negotiation or competition, and that there will not be any negative social consequences, the sexes are very similar.

Is competition good? To the extent that competition can help people do their best work, it is beneficial. Recall that our earlier review in chapter 2 of teams, where we considered the benefits of diversity, shows that cooperation is also good. Cooperation helps bring out everyone's talents. Institutions need to strike the right balance in order to maximize applications from a diverse pool of job candidates and to ensure that they bring out the best in all of their employees.

## 7. Overt Bias

Overt bias exists in academia, as the all too numerous documented cases of sexual harassment attest. For one example, women in geophysics, and especially women of color, reported disturbing levels of verbal and physical harassment in a 2015 Internet survey of men and women students, researchers, and faculty (Clancy, Lee, Rodgers, & Richey, 2017). Both men and women reported hearing sexist and racist comments, though women and people of color reported hearing them more than did White men. Students reported more instances than did more senior researchers. Since this is a sample of unknown representativeness, and included many more women than men, the data cannot inform us about how common harassment is, only that it occurs.[34] Even one occurrence is one too many. That harassment has consequences is clear from the larger percentage of women than men reporting feeling unsafe and missing professional events at higher levels because of feeling unsafe. Making matters worse is that academia has failed to prevent harassment or to adequately sanction harassers. Overt bias occurs and is damaging, even if there is now less of it and even if it is now more likely to meet with objections when it occurs. Professional societies and academic institutions have a duty to ensure that professional settings provide an environment where everyone is free from harassment or bullying. Some professional groups have taken steps to fulfill that duty, such as the American Geophysical Union, the American Political Science Association, and the Boston University Conference on Language Development. (See, respectively, https://harassment.agu.org/; http://www.apsanet.org/divresources/policyprocedures/; https://www.bu.edu/bucld/conference-info/conduct/.)

However, we argue that overt bias, which may be an extension of the more subtle examples we have described, is not the major determinant for the widespread pattern of sex, race, and ethnic disparities in achievement that we have documented. Even those women and people of color who never or seldom experience overt bias from peers or supervisors (the majority in the study of geophysicists; Clancy et al., 2017) experience the myriad types of subtle undervaluation that we have documented.

## Summary

To summarize, we have reviewed seven alternate accounts of the slower progress and lower salaries of women and people of color (demographic inertia, extra-academic responsibilities, ability, productivity, interest, negotiation and taste for competition, and overt discrimination):

- In the cases of negotiation, competition, and productivity, the facts the explanations are based on may themselves be due to unequal treatment of men and women. For example, men negotiate and compete more than women, but men who negotiate and compete are responded to more positively than women who negotiate and compete. Men are slightly more productive than women, but men and women have different access to resources.
- In the cases of ability and interest, there is insufficient evidence to conclude that men and women, Whites and people of color, differ. Here we conclude that the data to support the claims are not solid. In addition, like negotiation, competition, and productivity, skills, and tastes are affected by treatment.
- In the cases of demographic inertia and extra-academic responsibilities, the facts adduced are correct. Demographic inertia exists. But White men's success is greater than demographic inertia alone would predict. Extra-academic responsibilities differ by gender, and they remove women from the workforce, but they play a small role for those who remain.
- Finally, overt discrimination against White women and men and women of color exists and has negative consequences. It is responsible for some

portion of the slower advancement of women and people of color. But even women and people of color who do not experience overt discrimination advance more slowly than their White male peers.

We have repeatedly stated that most academics sincerely espouse egalitarian principles. We have also repeatedly stated that a belief in the merit principle makes it difficult for people to see the small ways in which they violate it. In later chapters, we recommend policies and procedures that will help people operate more consistently with their principles.

## Notes

1. We thank Francine Blau (personal communication, January 13, 2017) for bringing both points to our attention.

2. Data are from the National Science Board figure 3–31. Estimated salary differences between women and men with highest degree in S[cience]&E[ngineering] employed full-time, controlling for selected characteristics, by degree level: 2013. For BA/BS individuals, the disparity after all controls were entered was 8.4%; for MA/MS individuals, it was 7.8%. Access to the full report is here: https://nsf.gov/statistics/2016/nsb20161/#/report/

3. Data are from the National Science Board figure 3–32. Estimated salary differences between minorities and Whites and Asians with highest degree in S&E employed full-time, controlling for selected characteristics, by degree level, 2013. For BA/BS individuals, the disparity after all controls were entered was 4.9%; for MA/MS individuals, it was 5.5%. Access to the full report is here: https://nsf.gov/statistics/2016/nsb20161/#/report/

4. One study examining academic salaries by rank and subfield found few significant salary differences between men and women for 2010, though differences generally favored men (Ceci et al., 2014). Across the 24 field/rank combinations, 18 showed a numerical difference in favor of men; among research-intensive institutions, 16 of the 24 fields showed a numerical difference in favor of men. One difficulty with tabulating gender disparities separately by subfield and rank is that the reduction in sample size makes the data less stable; another is that the contribution of factors like ethnicity is unknown. There are also difficulties with having a large sample size, in particular that differences by field will be washed out. Why the NSF data, which controlled for field, show a disparity but the study by subfield does not show significant differences is not clear. Our own findings, reported below, also find differences by field.

5. Data are from table 12. Doctorate recipients, by major field of study: Selected years, 1985–2015. Retrieved online January 27, 2018, from https://www.nsf.gov

/statistics/2017/nsf17306/data/tab12.pdf. These are the most recent data available. If nonresident Asians were included, males would be the majority.

6. Calculations are based on NSF tables 19, 20, 21, and 24. Calculations in table 4.1 include only U.S. citizens and permanent residents. If nonresidents were included, the percentage of African Americans and "all others" would be reduced and the percentage of Asians would be increased; the percentage of women would be reduced. Different NSF tables have different types of missing data, resulting in discrepancies in totals from one table to another; the discrepancies are usually small. We include here the numbers and titles of the tables. Table 19: Doctorate recipients, by ethnicity, race, and citizenship status: 2005–15. Retrieved online January 27, 2018, from https://www.nsf.gov/statistics/2017/nsf17306/data/tab19.pdf. Total N for U.S. citizens and permanent nonresidents is 35,117. Table 20: Male doctorate recipients, by ethnicity, race, and citizenship status: 2005–15. Retrieved online January 27, 2018, from https://www.nsf.gov/statistics/2017/nsf17306/data/tab20.pdf. Total N for males who are U.S. citizens or permanent residents is 17,245. Table 21: Female doctorate recipients, by ethnicity, race, and citizenship status: 2005–15. Retrieved online January 27, 2018, from https://www.nsf.gov/statistics/2017/nsf17306/data/tab21.pdf. Total N for females who are U.S. citizens or permanent residents is 17,872.

7. Population data for 2014 were taken from NSF's table 1–2, Resident population of the United States, by sex, race or ethnicity, and age: 2014, retrieved January 27, 2018, from https://www.nsf.gov/statistics/2017/nsf17310/static/data/tab1-2.pdf. We included individuals between the ages of 25 and 64. "All others" include American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, two or more races, and other or unknown race and ethnicity (the largest subcategory).

8. We do not report statistical tests because (1) we do not have exact counts for the population and (2) some doctorate recipients are younger than age 25 and older than age 64. But the population numbers are so large that, even with the inexact counts, it is clear that were exact counts available, statistical tests would be significant for the comparisons. We computed some values using the population estimates for adults ages 25–64 with a Fisher exact probability test; even the smallest difference with the smallest overall numbers was significant.

9. The percentage of Asian men would be much higher if temporary visa holders were included.

10. Data are from table 14. Doctorate recipients, by broad field of study and sex: Selected years, 1985–2015. Retrieved January 27, 2018, from https://www.nsf.gov/statistics/2017/nsf17306/data/tab14.pdf. Note that this table includes all women, not simply women who were U.S. citizens or permanent residents.

11. Data are from table 23. U.S. citizen and permanent resident doctorate recipients, by broad field of study, ethnicity, and race: Selected years, 1995–2015. Retrieved January 27, 2018, from https://www.nsf.gov/statistics/2017/nsf17306/data/tab23.pdf.

12. Data are from table 9–21. Science, engineering, and health doctorate holders employed full-time in universities and four-year colleges, by 2005 Carnegie classification of academic institution, sex, race, ethnicity, and disability status: 2013; NSF; January 25, 2017, from http://www.nsf.gov/statistics/2015/nsf15311/tables/pdf/tab9-21.pdf. This table includes a column for missing data—individuals who did not report type of institution. We have subtracted the missing data from the totals that we use as the denominator. Depending on whether gender, ethnicity, or disability status is being examined, there are differing amounts of missing data.

13. Both gender differences are significant by a test of independent proportions. For very high research activity institutions, Males versus Females, $z=15.03$, $p<0.0002$; for high research activity institutions, Males versus Females, $z=16.27$, $p<0.0002$.

14. For Ethnicity, Whites are compared to every other group via a test of the significance of the difference between two independent proportions. At very high research activity institutions, Whites versus Asians, $z=-26.565$, $p<0.0002$ (Asians are better represented than Whites); Whites versus African Americans, $z=21.17$, $p<0.0002$; Whites versus Hispanics, $z=1.936$, $p=0.053$; Whites versus all others, $z=8.68$, $p<0.0002$. At high research activity institutions, Whites versus Asians, $z=-8.151$, $p<0.0002$ (Asians are better represented than Whites); Whites versus African Americans, $z=3.016$, $p<0.03$; Whites versus Hispanics, $z=4.091$, $p<0.0002$; Whites versus all others, $z=7.441$, $p<0.0002$.

We separately compared African Americans with every other group: they have significantly lower representation than any other group. At very high research activity institutions, African Americans versus Asians, $z=-31.836$, $p<0.0002$; African Americans versus Hispanics, $z=-15.12$, $p<0.0002$; African Americans versus all others, $z=-5.777$, $p<0.0002$. At high research activity institutions, African Americans versus Asians, $z=-6.405$, $p<0.0002$; African Americans versus Hispanics, $z=.5$, ns; African Americans versus all others, $z=4.569$, $p<0.0002$.

15. Disability differences are significant by a test of the significance of the difference between independent proportions. At very high research activity institutions, individuals without disabilities are better represented than those with disabilities, $z=9.429$, $p<0.0002$; at high research activity institutions, individuals without disabilities are worse represented than those with disabilities, $z=-6.17$, $p<0.0002$.

16. Data are from table 18. Employed doctoral scientists and engineers in four-year educational institutions, by broad field of doctorate, sex, faculty rank, and years since doctorate: 2013. Retrieved January 27, 2018, from https://ncsesdata.nsf.gov/doctoratework/2013/html/SDR2013_DST18.html. Note that NSF rounds all data to the nearest hundred, resulting in some small discrepancies. In all cases we accepted the values in individual cells and computed subtotals as required.

17. NSF does not provide similar data by race and ethnicity, nor does it present U.S. citizens separately. Although NSF does present data for subfields within science, we

do not look at subfields because several cells are empty because of the small number of people who reach full professor status within ten years.

18. Tests of independent proportions revealed significantly more men than women for the combined ranks of full and associate professors for *all fields*, $z = 10.00$, $p < 0.0002$ and for *science*, $z = 13.32$, $p < 0.0002$; and more women than men for *health*, $z = -2.69$, $p < 0.01$. A test of independent proportions for the associate professor level for *engineering* showed significantly more men than women, $z = 6.01$, $p < 0.0002$.

19. When all individuals in *all fields* are included (N = 101,500) 16% of men and 13% of women are in the top two ranks, a difference which is significant by a test of independent proportions, $z = 11.88$, $p < 0.0002$. For all individuals in *science* fields (N = 80,600), 15% of men and 11% of women are in the top two ranks, a difference which is significant by a test of independent proportions, $z = 16.12$, $p < 0.0002$. The same holds for tenure proportions, presented later. (It is not clear why the totals for tenured plus tenure track individuals are smaller than the totals for full, associate, plus assistant professors.)

20. We conducted separate tests of independent proportions to determine whether there was a significant difference in the percentage of men and women at the associate professor rank. Men are better represented than women at the associate professor rank in *mathematics and statistics*, $z = 11.62$, $p < 0.0002$; the *physical sciences*, $z = 2.043$, $p < .05$; the *biological and agricultural sciences*, $z = 10.49$, $p < 0.0002$; *social science (excluding psychology)*, $z = 4.82$, $p < 0.0002$.

21. Tests of independent proportions showed no difference between men and women in their proportion of achieving the rank of associate professor within ten years in *computer science*, $z = 1.77$, ns, or *psychology*, where the percentages of men and women achieving associate professor rank are identical.

22. See table 21. Employed doctoral scientists and engineers in four-year educational institutions, by broad field of doctorate, sex, tenure status, and years since doctorate: 2013. Retrieved January 27, 2018, from https://ncsesdata.nsf.gov/doctoratework/2013/html/SDR2013_DST21.html. Some totals here differ from totals in table 18. There are a few small rounding inconsistencies in NSF's table 21. In all cases we accepted the values in individual cells and computed subtotals as required.

23. Tests of independent proportions revealed that, for *all fields*, men are significantly more likely than women to achieve tenure within ten years, $z = 2.90$; $p < 0.004$. The same holds for *science as a whole*, $z = 6.57$; $p < 0.0002$. In *engineering* there is no difference, $z = -0.81$, ns. For *health*, women are more likely than men to achieve tenure within ten years, $z = -5.28$, $p = 0.0002$.

24. We computed individual tests of independent proportions for achievement of tenure within ten years for each field. For *computer/information sciences*, there was no gender difference, $z = 1.77$, ns, mirroring the results for achievement of associate

professor rank in ten years. In contrast, for *mathematics/statistics,* men are more likely to achieve tenure, $z=9.21$, $p<0.0002$. The same is true for *physical sciences*, $z=6.04$, $p<0.0002$, and *biological/agricultural/environmental life sciences*, $z=8.07$, $p<0.0002$. There are no gender differences for *social sciences*, $z=-0.599$, ns. In *psychology*, women are more likely than men to achieve tenure within ten years, $z=-2.86$, $p<0.005$.

25. Ceci et al. (2014) adopt a similar strategy of considering different explanations in their case for the underrepresentation of women in the sciences. Some of the explanations they consider overlap with those we consider here. They sensibly conclude that none of the explanations they consider will, on their own, account for women's lower representation. Blau and Kahn (2017) similarly consider a range of different accounts for gender disparities in the workforce as a whole, concluding that measurable human and social capital factors are not sufficient to explain all of the disparities.

26. The pipeline metaphor may incorrectly make it seem as if there is only one route to science and may be demotivating. For discussion, see Cannady, Greenwald, and Harris (2014); Miller and Wai (2015); and Xie and Shauman (2003).

27. As academic salaries become ever more compressed and rise ever more slowly, men might leave in greater numbers than women for more lucrative fields. If women's salaries are depressed no matter what economic sector they are in, they will have less reason to desert academia.

28. If one assumes, as Shaw and Stanton (2012) do, that movement into a tenured position requires that older individuals leave that position, demographic inertia does account for the transition from untenured to tenured faculty member. But since that simplifying assumption is incorrect, the model using it cannot appropriately test the demographic inertia hypothesis for tenure status.

29. In 1981–1985, seventh graders who took the math SAT as a possible entry point to a program for gifted students showed a gender disparity in favor of boys among the top 1% (1.43 times as many boys as girls), top 0.5% (2.61 times as many boys as girls), and top 0.01% (13.5 times as many boys as girls). For the next five-year period, 1986–1990, those ratios were reduced to 1.31, 2.15, and 7.6. For 1991–1995, they were reduced again to 1.26, 1.95, and 3.87, and for 2006–2010, the data showed lower ratios for the 1% (1.10) and 0.5% (1.54) levels and a similar ratio at the 0.01% level (3.83; Wai et al., 2010). The 2011–2015 cohort showed another reduction in the sex ratio at the uppermost end, to 2.53:1 (Makel et al., 2016). ACT math scores have also shown a decline in the ratio at the 0.01% level, from 3.14 in the 1990–1995 cohort to 2.6 in the 2006–2010 cohort. There may be differences in how the SAT and ACT are constructed, or in the percentages of boys and girls who take the tests, since the male:female ratios among top scorers are different.

Two things are notable. One is the rapid increase in the percentage of girls in the far right tail of the SAT distribution. That increase could be an artifact, in that a different sample of girls may have taken the test in the first decade after the 1981–1985 cohort, or the test could have changed (though Wai et al., 2010, argue that the changes are not meaningful). If it is not an artifact, the difference in the male:female ratio for the SAT shows the lability of math scores, as does the ACT reduction to a lesser extent. Apparently, social context influences children's math performance and can rapidly reduce a 13.5 ratio to a 2.53 ratio, even though the reduction is not linear. The same reduction in the sex ratios was found in India (Makel et al., 2016).

30. There are average differences among different racial and ethnic groups in standardized verbal and math scores. There is evidence both that the tests are less valid predictors of performance for African Americans and Hispanics than they are for Whites (Berry, Clark, & McClure, 2011; Berry, Cullen, & Meyer, 2014), and that cognitive tests slightly overpredict how well African Americans and Hispanics will perform (see Berry & Zhao, 2015; for reviews, see Berry, 2015, and Schmitt, 2014).

31. Math and verbal scores are highly intercorrelated; for the sample of gifted 13-year-olds the correlation is 0.55 (Lubinski, Webb, Morelock, & Benbow, 2001). Girls are more likely than boys to have high verbal scores. That asymmetry—girls can succeed in more fields than boys can—might play a role in field choice.

32. There are a number of different measures of citations, such as total number, average number, *h*, and so on.

33. There are a number of technical issues, such as whether means, medians, or distributions should be used in comparing men and women. Male authors tend to dominate the very highest numbers of publications, and women tend to dominate the lowest numbers, so means may overestimate differences. Researchers and scholars at research-intensive universities publish more than do their peers at four-year colleges, where women are more highly represented. Older studies in general show a greater disparity in favor of men; more recent studies show less disparity.

34. One analysis (Greco, O'Boyle, & Walter, 2015) suggests that studies of counterproductive work behavior have lower response rates than other survey studies and are likely to underreport the existence of troubling rates of inappropriate comments, harassment, and the like.

## References

Ahmad, S. (2016). Family or future in the academy? *Review of Educational Research, 87*(1), 204–239. doi:10.3102/0034654316631626

Amanatullah, E. T., & Tinsley, C. H. (2013). Ask and ye shall receive? How gender and status moderate negotiation success. *Negotiation and Conflict Management Research*, *6*(4), 253–272.

Andersen, S., Ertac, S., Gneezy, U., List, J. A., & Maximiano, S. (2013). Gender, competitiveness, and socialization at a young age: Evidence from a matrilineal and a patriarchal society. *Review of Economics and Statistics*, *95*(4), 1438–1443.

Anghel, B., de la Rica, S., & Dolado, J. J. (2011). The role of institutions in gender differences in science careers. In Ministerio de Ciencia e Innovación (Ed.), *White paper on the position of women in science in Spain*. Retrieved January 27, 2018, from http://www.idi.mineco.gob.es/stfls/MICINN/Ministerio/FICHEROS/UMYC/WhitePaper_Interactive.pdf

Ayres, I., Banaji, M., & Jolls, C. (2015). Race effects on eBay. *Rand Journal of Economics*, *46*(4), 891–917.

Ayres, I., & Siegelman, P. (1995). Race and gender discrimination in bargaining for a new car. *American Economic Review*, *85*(3), 304–321.

Babcock, L., & Laschever, S. (2009). *Women don't ask: Negotiation and the gender divide*. Princeton, NJ: Princeton University Press.

Benard, S., & Correll, S. J. (2010). Normative discrimination and the motherhood penalty. *Gender & Society*, *24*(5), 616–646.

Bentley, P. J., & Kyvik, S. (2013). Individual differences in faculty research time allocations across 13 countries. *Research in Higher Education*, *54*(3), 329–348.

Ben-Zeev, T., Fein, S., & Inzlicht, M. (2005). Arousal and stereotype threat. *Journal of Experimental Social Psychology*, *41*(2), 174–181.

Berry, C. M. (2015). Differential validity and differential prediction of cognitive ability tests: Understanding test bias in the employment context. *Annual Review of Organizational Psychology and Organizational Behavior*, *2*(1), 435–463.

Berry, C. M., Clark, M. A., & McClure, T. K. (2011). Racial/ethnic differences in the criterion-related validity of cognitive ability tests: A qualitative and quantitative review. *Journal of Applied Psychology*, *96*, 881–906. doi:10.1037/a0023222

Berry, C. M., Cullen, M. J., & Meyer, J. M. (2014). Racial/ethnic subgroup differences in cognitive ability test range restriction: Implications for differential validity. *Journal of Applied Psychology*, *99*(1), 21–37.

Berry, C. M., & Zhao, P. (2015). Addressing criticisms of existing predictive bias research: Cognitive ability test scores still overpredict African Americans' job performance. *Journal of Applied Psychology*, *100*(1), 162–179.

Bertrand, M., Goldin, C., & Katz, L. F. (2010). Dynamics of the gender gap for young professionals in the financial and corporate sectors. *American Economic Journal. Applied Economics*, *2*(3), 228–255.

Bianchi, S. M., Sayer, L. C., Milkie, M. A., & Robinson, J. P. (2012). Housework: Who did, does, or will do it, and how much does it matter? *Social Forces*, *91*(1), 55–63.

Blau, F. D., & Kahn, L. M. (2017). The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature*, *55*, 789–865.

Bombardieri, M. (April 8, 2005). Summers displays new understanding of women's careers. *The Boston Globe*. http://archive.boston.com/news/education/higher/articles/2005/04/08/summers_displays_new_understanding_of_womens_careers/

Boring, A. (2017). Gender biases in student evaluations of teaching. *Journal of Public Economics*, *145*, 27–41.

Boring, A., Ottoboni, K., & Stark, P. (2016). Student evaluations of teaching (mostly) do not measure teaching effectiveness. *ScienceOpen Research*, 1–11. https://www.scienceopen.com/document/read?vid=818d8ec0-5908-47d8-86b4-5dc38f04b23e

Bowles, H. R., Babcock, L., & Lai, L. (2007). Social incentives for gender differences in the propensity to initiate negotiations: Sometimes it does hurt to ask. *Organizational Behavior and Human Decision Processes*, *103*(1), 84–103.

Box-Steffensmeier, J. M., Cunha, R. C., Varbanov, R. A., Hoh, Y. S., Knisley, M. L., & Holmes, M. A. (2015). Survival analysis of faculty retention and promotion in the social sciences by gender. *PLoS One*, *10*(11), e0143093.

Byrnes, J. P. (2007). Publishing trends of psychology faculty during their pretenure years. *Psychological Science*, *18*, 283–286.

Cameron, E. Z., White, A. M., & Gray, M. E. (2016). Solving the productivity and impact puzzle: Do men outperform women, or are metrics biased? *Bioscience*, *66*(3), 245–252. doi:10.1093/biosci/biv173

Cannady, M. A., Greenwald, E., & Harris, K. N. (2014). Problematizing the STEM pipeline metaphor: Is the STEM pipeline metaphor serving our students and the STEM workforce? *Science Education, 98*, 443–460. doi:10.1002/sce.21108

Ceci, S. J., Ginther, D. K., Kahn, S., & Williams, W. M. (2014). Women in academic science: A changing landscape. *Psychological Science in the Public Interest*, *15*(3), 75–141.

Ceci, S. J., & Williams, W. M. (2010). Sex differences in math-intensive fields. *Current Directions in Psychological Science*, *19*(5), 275–279.

Cheryan, S., Master, A., & Meltzoff, A. N. (2015). Cultural stereotypes as gatekeepers: Increasing girls' interest in computer science and engineering by diversifying stereotypes. *Frontiers in Psychology*, *6*, article 49. doi:10.3389/fpsyg.2015.00049

Cheryan, S., & Plaut, V. C. (2010). Explaining underrepresentation: A theory of precluded interest. *Sex Roles*, *63*(7–8), 475–488.

Cheryan, S., Plaut, V. C., Davies, P. G., & Steele, C. M. (2009). Ambient belonging: How stereotypical cues impact gender participation in computer science. *Journal of Personality and Social Psychology*, *97*(6), 1045–1060.

Cheryan, S., Ziegler, S. A., & Jiang, L. (2017). Why are some STEM fields more gender balanced than others? *Psychological Bulletin*, *143*, 1–35.

Clancy, K. B. H., Lee, K. M. N., Rodgers, E. M., & Richey, C. (2017). Double jeopardy in astronomy and planetary science: Women of color face greater risks of gendered and racial harassment. *Journal of Geophysical Research. Planets*, *122*. doi:10.1002/2017JE005256

Correll, S. J., Benard, S., & Paik, I. (2007). Getting a job: Is there a motherhood penalty? *American Journal of Sociology*, *112*(5), 1297–1338.

Croson, R., & Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, *47*(2), 448–474.

Dasgupta, N., & Stout, J. G. (2014). Girls and women in science, technology, engineering, and mathematics: STEMing the tide and broadening participation in STEM careers. *Policy Insights from the Behavioral and Brain Sciences*, *1*(1), 21–29.

Diekman, A. B., Brown, E. R., Johnston, A. M., & Clark, E. K. (2010). Seeking congruity between goals and roles: A new look at why women opt out of science, technology, engineering, and mathematics careers. *Psychological Science*, *21*, 1051–1057.

Doleac, J. L., & Stein, L. C. (2013). The visible hand: Race and online market outcomes. *Economic Journal (London)*, *123*(572), F469–F492.

Duch, J., Zeng, X. H. T., Sales-Pardo, M., Radicchi, F., Otis, S., Woodruff, T. K., & Amaral, L. A. N. (2012). The possible role of resource requirements and academic career-choice risk on gender differences in publication rate and impact. *PLoS One*, *7*(12), e51332. doi:10.1371/annotation/7f54a3e6-6dcf-4825-9eb9-201253cf1e25

Ellison, G., & Swanson, A. (2010). The gender gap in secondary school mathematics at high achievement levels: Evidence from the American Mathematics Competitions. *Journal of Economic Perspectives*, *24*(2), 109–128.

Flory, J. A., Leibbrandt, A., & List, J. A. (2015). Do competitive workplaces deter female workers? A large-scale natural field experiment on job-entry decisions. *Review of Economic Studies*, *82*, 122–155.

Fox, M. F. (2005). Gender, family characteristics, and publication productivity among scientists. *Social Studies of Science*, *35*(1), 131–150.

Ginther, D. K., & Kahn, S. (November 2006). Does science promote women? Evidence from academia 1973–2001 (NBER Working Paper 12691). Cambridge, MA: National Bureau of Economic Research. http://www.nber.org/papers/w12691

Ginther, D. K., Kahn, S., & Schaffer, W. T. (2016). Gender, race/ethnicity, and National Institutes of Health R01 research awards: Is there evidence of a double bind for women of color? *Academic Medicine*, *91*(8), 1098–1107.

Ginther, D. K., Schaffer, W. T., Schnell, J., Masimore, B., Liu, F., Haak, L. L., & Kington, R. (2011). Race, ethnicity, and NIH research awards. *Science*, *333*(6045), 1015–1019.

Glynn, A. N., & Sen, M. (2015). Identifying judicial empathy: Does having daughters cause judges to rule for women's issues? *American Journal of Political Science*, *59*(1), 37–54.

Gneezy, U., Leonard, K. L., & List, J. A. (2009). Gender differences in competition: Evidence from a matrilineal and a patriarchal society. *Econometrica*, *77*(5), 1637–1664.

Goulden, M., Mason, M. A., & Frasch, K. (2011). Keeping women in the science pipeline. *Annals of the American Academy of Political and Social Science*, *638*(1), 141–162.

Greco, L. M., O'Boyle, E. H., & Walter, S. L. (2015). Absence of malice: A meta-analysis of nonresponse bias in counterproductive work behavior research. *Journal of Applied Psychology*, *100*(1), 75–97. doi:10.1037/a0037495

Hargens, L. L., & Long, J. S. (2002). Demographic inertia and women's representation among faculty in higher education. *Journal of Higher Education*, *73*, 494–517.

Heilman, M. E. (2012). Gender stereotypes and workplace bias. *Research in Organizational Behavior*, *32*, 113–135.

Heilman, M. E., & Okimoto, T. G. (2007). Why are women penalized for success at male tasks? The implied communality deficit. *Journal of Applied Psychology*, *92*(1), 81–92.

Iacus, S. M., King, G., & Porro, G. (2011). Multivariate matching methods that are monotonic imbalance bounding. *Journal of the American Statistical Association*, *106*(493), 345–361.

Jagsi, R., Griffith, M. K. A., Stewart, A., Sambuco, M. D., DeCastro, M. R., & Ubel, P. A. (2013). Gender differences in salary in a recent cohort of early-career physician-researchers. *Academic Medicine*, *88*(11), 1689–1699.

Joy, S. (2006). What should I be doing, and where are they doing it? Scholarly productivity of academic psychologists. *Perspectives on Psychological Science*, *1*, 346–364.

Kahn, S., & Ginther, D. (2017). Women and STEM (NBER Working Paper 23525). Cambridge, MA: National Bureau of Economic Research.

Kaminski, D., & Geisler, C. (2012). Survival analysis of faculty retention in science and engineering by gender. *Science*, *335*(6070), 864–866.

Kelly, K., & Grant, L. (2012). Penalties and premiums: The impact of gender, marriage, and parenthood on faculty salaries in science, engineering and mathematics (SEM) and non-SEM fields. *Social Studies of Science*, *42*(6), 869–896.

King, E. G., Botsford, W., Hebl, M. R., Kazama, S., Dawson, J. F., & Perkins, A. (2012). Benevolent sexism at work: Gender differences in the distribution of challenging developmental experiences. *Journal of Management*, *38*, 1835–1866.

Kray, L. J., & Thompson, L. (2004). Gender stereotypes and negotiation performance: An examination of theory and research. *Research in Organizational Behavior*, *26*, 103–182.

Kulik, C. T., & Olekalns, M. (2012). Negotiating the gender divide lessons from the negotiation and organizational behavior literatures. *Journal of Management*, *38*(4), 1387–1415.

Larivière, V., Ni, C., Gingras, Y., Cronin, B., & Sugimoto, C. R. (2013). Global gender disparities in science. *Nature*, *504*, 211–213.

Leahey, E. (2006). Gender differences in productivity—Research specialization as a missing link. *Gender & Society*, *20*(6), 754–780.

Leibbrandt, A., & List, J. A. (2014). Do women avoid salary negotiations? Evidence from a large-scale natural field experiment. *Management Science*, *61*(9), 2016–2024.

Lubinski, D., Benbow, C. P., & Kell, H. J. (2014). Life paths and accomplishments of mathematically precocious males and females four decades later. *Psychological Science*, *25*(12), 2217–2232.

Lubinski, D., Webb, R. M., Morelock, M. J., & Benbow, C. P. (2001). Top 1 in 10,000: A 10-year follow-up of the profoundly gifted. *Journal of Applied Psychology*, *86*(4), 718–729.

MacNell, L., Driscoll, A., & Hunt, A. N. (2015). What's in a name: Exposing gender bias in student ratings of teaching. *Innovative Higher Education*, *40*(4), 291–303.

Makel, M. C., Wai, J., Peairs, K., & Putallaz, M. (2016). Sex differences in the right tail of cognitive abilities: An update and cross cultural extension. *Intelligence*, *59*, 8–15.

Manning, A., & Swaffield, J. (2008). The gender gap in early-career wage growth. *Economic Journal (London)*, *118*(530), 983–1024.

Mason, M. A., & Goulden, M. (2004). Marriage and baby blues: Redefining gender equity in the academy. *Annals of the American Academy of Political and Social Science*, *596*(1), 86–103.

Mason, M. A., Wolfinger, N. H., & Goulden, M. (2013). *Do babies matter? Gender and family in the ivory tower*. New Brunswick, NJ: Rutgers University Press.

Mazei, J., Hüffmeier, J., Freund, P. A., Stuhlmacher, A. F., Bilke, L., & Hertel, G. (2015). A meta-analysis on gender differences in negotiation outcomes and their moderators. *Psychological Bulletin*, *141*(1), 85–104.

Melguizo, T., & Strober, M. H. (2007). Faculty salaries and the maximization of prestige. *Research in Higher Education*, *48*(6), 633–668.

Miller, D. I., & Wai, J. (2015). The bachelor's to Ph.D. STEM pipeline no longer leaks more women than men: A 30-year analysis. *Frontiers in Psychology, 6*, article 37. doi:10.3389/fpsyg.2015.00037

Mitchell, S. M., & Hesli, V. L. (2013). Women don't ask? Women don't say no? Bargaining and service in the political science profession. *PS, Political Science & Politics*, *46*(02), 355–369.

National Science Board. (2016). *Science and engineering indicators 2016* (pp. NSB-2016–NSB-1). Arlington, VA: National Science Foundation.

Niederle, M. (2015). Gender. In J. H. Kagel & A. E. Roth (Eds.), *The handbook of experimental economics* (Vol. *2*, pp. 481–562). Princeton, NJ: Princeton University Press.

Niederle, M., & Vesterlund, L. (2010). Explaining the gender gap in math test scores: The role of competition. *Journal of Economic Perspectives*, *24*(2), 129–144.

Noftle, E. E., & Robins, R. W. (2007). Personality predictors of academic outcomes: Big five correlates of GPA and SAT scores. *Journal of Personality and Social Psychology*, *93*(1), 116–130.

O'Meara, K., Kuvaeva, A., & Nyunt, G. (2017). Constrained choices: A view of campus service inequality from annual faculty reports. *Journal of Higher Education*, *88*(5), 672–700. doi:10.1080/00221546.2016.1257312

O'Meara, K., Kuvaeva, A., Nyunt, G., Waugaman, C., & Jackson, R. (in press). Asked more often: Gender differences in faculty workload in research universities and the work interactions that shape them. *American Educational Research Journal*.

O'Meara, K., Sandmann, L. R., Saltmarsh, J., & Giles, D. E. (2011). Studying the professional lives and work of faculty involved in community engagement. *Innovative Higher Education*, *36*(2), 83–96.

Parks, A. L., & Redberg, R. F. (2017). Women in medicine and patient outcomes: Equal rights for better work? *JAMA Internal Medicine*, *177*, 161.

Ridgeway, C. L. (1982). Status in groups: The importance of motivation. *American Sociological Review*, *47*, 175–188.

Ridgeway, C. L. (2011). *Framed by gender: How gender inequality persists in the modern world*. Oxford, UK: Oxford University Press.

Rudman, L. A., & Fairchild, K. (2004). Reactions to counterstereotypic behavior: The role of backlash in cultural stereotype maintenance. *Journal of Personality and Social Psychology*, *87*, 157–176.

Rudman, L. A., & Phelan, J. E. (2008). Backlash effects for disconfirming gender stereotypes in organizations. *Research in Organizational Behavior*, *28*, 61–79.

Schmidt, F. L. (2011). A theory of sex differences in technical aptitude and some supporting evidence. *Perspectives on Psychological Science*, *6*, 560–573.

Schmitt, N. (2014). Personality and cognitive ability as predictors of effective performance at work. *Annual Review of Organizational Psychology and Organizational Behavior*, *1*(1), 45–65.

Schroen, A. T., Brownstein, M. R., & Sheldon, G. F. (2004). Women in academic general surgery. *Academic Medicine*, *79*(4), 310–318.

Shaw, A. K., & Stanton, D. E. (2012). Leaks in the pipeline: Separating demographic inertia from ongoing gender differences in academia. *Proceedings. Biological Sciences*, *279*, 3736–3741.

Shurchkov, O. (2012). Under pressure: Gender differences in output quality and quantity under competition and time constraints. *Journal of the European Economic Association*, *10*, 1189–1213.

Sio, U. N., & Ormerod, T. C. (2009). Does incubation enhance problem solving? A meta-analytic review. *Psychological Bulletin*, *135*(1), 94–120.

Smith, J. L., Lewis, K. L., Hawthorne, L., & Hodges, S. D. (2013). When trying hard isn't natural: Women's belonging with and motivation for male-dominated STEM fields as a function of effort expenditure concerns. *Personality and Social Psychology Bulletin*, *39*(2), 131–143.

Sutanto, W., Bell, R. L., Fei, Q., & Scott, J. (2014). Is there a gender pay gap in business schools? *Business Studies Journal*, *6*(2), 39–56.

Thomas, N. R., Poole, D. J., & Herbers, J. M. (2015). Gender in science and engineering faculties: Demographic inertia revisited. *PLoS One*, *10*(10), e0139767.

Tsugawa, Y., Jena, A. B., Figueroa, J. F., Orav, E. J., Blumenthal, D. M., & Jha, A. K. (2017). Comparison of hospital mortality and readmission rates for Medicare patients treated by male vs female physicians. *JAMA Internal Medicine*, *177*, 206–213.

Valian, V. (2014). Interests, gender, and science. *Perspectives on Psychological Science*, *9*(2), 225–230.

van den Besselaar, P., & Sandström, U. (2016). Gender differences in research performance and its impact on careers: A longitudinal case study. *Scientometrics*, *106*(1), 143–162.

Wagner, N., Rieger, M., & Voorvelt, K. (2016). Gender, ethnicity and teaching evaluations: Evidence from mixed teaching teams. *Economics of Education Review*, *54*, 79–94.

Wai, J., Cacchio, M., Putallaz, M., & Makel, M. C. (2010). Sex differences in the right tail of cognitive abilities: A 30 year examination. *Intelligence*, *38*(4), 412–423.

Washington, E. L. (2008). Female socialization: How daughters affect their legislator fathers' voting on woman's issues. *American Economic Review*, *98*(1), 311–332.

Weeden, K. A., Thébaud, S., & Gelbgiser, D. (2017). Degrees of difference: Gender segregation of US doctorates by field and program prestige. *Sociological Science*, *4*, 123–150.

Williams, W. M., & Ceci, S. J. (2015). National hiring experiments reveal 2:1 faculty preference for women on STEM tenure track. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(17), 5360–5365.

Xie, Y., & Shauman, K. A. (2003). *Women in science: Career processes and outcomes*. Cambridge, MA: Harvard University Press.