A Lay-Statistician Explanation of Minority Discrimination: A Research Note

Brent Simpson (University of South Carolina)*

and

Kimmo Eriksson (Mälardalen University and Stockholm University)

August 2011

* Direct correspondence to Brent Simpson, Department of Sociology, University of South Carolina, Columbia SC 29208. Email: bts@sc.edu.

A Lay-Statistician Explanation of Minority Discrimination: A Research Note

We outline a new explanation of discrimination against numerical minorities. In contrast to prior work that focuses on how the *content* of categories affects discrimination, our arguments describes how the *size* of categories discrimination. Specifically, we argue that, when comparing multiple categories, actors tend to view larger categories as more representative of an underlying population than smaller ones. As a result, a decision maker will tend to expect that members of a numerical majority are more likely to be what s/he is searching for, whether it is the best *or worst* candidate. We report the results of two studies designed to test these arguments. To demonstrate the generality of the proposed mechanism, Study 1 tested the argument in a non-social domain. Participants disproportionately favored the majority (versus minority) category when searching for a single winning lottery ticket, and favored the minority category when the goal is to avoid a single losing ticket. Our second study supported an additional implication of the argument in a social domain: decision makers tended to rank highly qualified majority job candidates as better than equally qualified minority candidates, and relatively unqualified majority candidates as worse than equally unqualified minority candidates.

A Lay Statistician Explanation of Minority Discrimination: A Research Note

Although, women and minorities have made important strides in the labor market over the past several decades, they still face discrimination (Tomaskovic-Devey et al. 2006; Pager and Shepherd, 2008; Ridgeway 2011). For instance, they are less likely to be hired for or promoted to prestigious positions, compared to their male and white counterparts (Catalyst 1996; Reskin 2001; National Research Council 2004; Correll & Benard 2006). Similarly, research shows high levels of sex- and race-segregation, with females and minorities disproportionately represented in sectors or occupations with lower pay (Reskin 1993; Tomaskovic-Devey et al. 2006). From dominant theoretical perspectives, this persistent discrimination is somewhat surprising, given growing evidence that gender- and racial attitudes have become more egalitarian (e.g., Dovidio & Gaertner 2004; Bobo 2001; Foschi & LaPointe 2002; Pager and Shepherd, 2008), as well as the introduction of federal legislation barring discrimination based on social category membership (Burstein and Edwards 1994; Quillian 2006). But, as Reskin (2002) notes, legislation and policies aimed at barring discrimination tend to assume that discrimination is based on decision-makers' intentional or rational choices. As such, these policies ignore "an even more important reason individuals' sex and race are routinely and illegitimately linked to employment rewards: automatic nonconscious cognitive processes that distort our perceptions and treatment of others." Reskin (2002) and others (e.g., National Research Council 2004; Quillian 2006) have called for investigation into these biases and how they produce discrimination in the labor market.

Here we build on the rich tradition of research on cognitive biases and heuristics (Kahneman, Slovic, & Tversky 1982; Gigerenzer and Gassmaier 2011) to outline and test a novel "lay-statistician" approach to discrimination. Our argument is based on two assumptions, explained in detail below: decision-makers *i*) categorize others (Macrae & Bodenhausen 2000) and *ii*) view larger samples or categories as more likely than smaller ones to contain whomever they are searching for (Evans & Dusoir 1977). As a result, a decision maker will tend to expect that the best candidate will be in the majority category. As we show below, our argument also predicts that decision-makers will expect that the *worst* candidate is most likely to be in the majority category. This leads to a preference for members of numerical majorities when selecting the best candidate and a preference for members of numerical minorities when attempting to avoid the worst candidate.

We develop this account more fully in the two sections to follow, and then present two tests of the arguments. To demonstrate that the hypothesized process reflects a general feature of how people think about categories, our first study addresses a non-social case (strategies for choosing among lottery tickets of two different colors). Our second study involves social categories (ranking job applicants about whom category information is available).

Social Categories and Discrimination

When making decisions about whom to hire or promote, employers often must choose between applicants who differ from each other along a number of dimensions. Some of these dimensions may be job-relevant (e.g., education and prior work experience). But social categories, such as applicants' race- or gender-categories, also have powerful effects on hiring decisions. Previous work has pointed to a number of reasons that social categories like race and gender impact hiring decisions (for a review, see National Research Council 2004). Broadly, as suggested by Reskin (2002), most traditional explanations of category-based discrimination have been on animosity or negative feelings toward a given category (e.g., Allport 1954; Becker 1957) or "rational" responses to the uncertainty facing employers. In these latter, "statistical discrimination" approaches (Arrow 1973; Aigner and Cain 1977; Lundberg and Startz 1983; England 1992), an employer has limited information about a prospective employee's skills or productivity for a given employee. To cope with this uncertainty, the employer may favor candidates from one category (whites, or men) over another (African Americans, or women). This preference can be rational if the employer believes with some degree of certainty that there are differences in the distribution of skills and talents in one category vs. another.[1]

While these dominant explanations offer important insights into bases of inequality in hiring and promotion, there is growing evidence that much discrimination is based on non-rational and "automatic" biases (Reskin 2002; National Resource Council 2004). Most of this work addresses how social categories influence evaluations, often outside conscious awareness. For instance, experimental studies show that knowledge of another's race-category can generate stereotypical cognitions (Bodenhausen et al. 1998) and associated expectations (Fiske 1998). Moreover, subliminally priming race can lead to stereotype-consistent behaviors. For instance, a study by Bargh et al. (1996) showed that white participants primed with black faces subsequently behaved in a more hostile way than did those primed with white faces.(See Reskin [2002] for a review and discussion of how these implicit and automatic processes generate category-based discrimination in labor market contexts.)

Our goal is to contribute to this growing literature by point to a heretofore unidentified process through which category-based discrimination occurs. Consistent with the literature

[1] Different models of statistical discrimination make different assumptions about whether the decision-maker's beliefs correspond to *actual* category differences in distributions of skills and productivity (see Correll and Benard 2006, for a review and comparison of existing models.)

reviewed thus far, the first basic assumption underlying our argument is that decision-makers think categorically. As Allport (1954, 19) put it, humans "cannot possibly avoid" thinking in terms of social categories. Upon meeting another person, we automatically encode the person's race and gender, in addition to her age (cite Kurzban et al. XXXX; cites from Kurzban; Fiske 1998). Moreover, an extensive tradition of research on "minimal groups" demonstrates the tendency for humans to attend to, and discriminate on the basis of, even arbitrary categorical distinctions (SIT cites; Dale Miller). Due to the tendency for people to treat irrelevant categories *as if* they are meaningful and relevant, categorical distinctions often form the basis of ingroup-favoritism (cites) and consensual status beliefs (Ridgeway 1991; Ridgeway et al. 1998). In short, decision-makers categorize others based on characteristics such as race and gender, and these resulting categories influence perceptions of -- and behavior toward -- category members in a range of ways. As such, as Reskin (2002, p. 221) noted, categorization is the "core cognitive process that links race and gender to workplace discrimination."

Here we argue that the tendency to think in categorical terms can have important and heretofore unrecognized effects on hiring and promotion decisions. As explained in the section to follow, all other things equal, decision makers will tend to think that they will be most likely to find whatever they are looking for, such as the best candidate, in the larger category. As a result, they will tend to give preferential treatment to majority candidates when attempting to select the best candidate. This is problematic when the basis of categorization is arbitrary or category membership is orthogonal to job performance.

Intuitive Statisticians and the Law of Large Numbers

A key question in the research reviewed thus far is how the *content* of categories impacts beliefs about – and treatment of - members of different categories. We are interested instead in how the *size* of categories impacts treatment of category members. That is, we know from research reviewed in the prior section that employers will tend to categorize job candidates based on characteristics such as race and gender. We address how the relative sizes of the resulting categories leads to discrimination against members of the numerical minority.

Key to our *lay-statistician* theory of discrimination is that unequal category sizes can lead to preferential treatment of one category versus another. Importantly, employers routinely face such situations, especially when searching for candidates for higher-paying jobs. For a number of reasons, including the basic fact that there are fewer minorities in the population, these jobs often have very few racial and ethnic minority applicants (cites). Similarly, prior work shows that females apply in smaller numbers to high paying jobs (Fernandez and Mors 2008; Fernandez and Abraham 2010). As a result, those seeking to fill high paying and prestigious positions are often faced with a relatively large number of white and male applicants. We suggest that the disparity in relative sizes of the applicant pools has an important, and heretofore over-looked, effect on preferences toward numerical majorities vs. minorities: all other things equal, an employer will tend to assume that he or she will be most apt to find the best candidate in the numerical majority category.

Following prior work (e.g., Kanter 1977; Salancik and Pfeffer 1978; Reskin 2002; National Research Council 2004), we suggest that decision-makers will be influenced by categorical information under conditions of uncertainty. That is, because employers typically lack direct behavioral evidence for how prospective employers will perform, decisions about whom to hire generally entail high levels of uncertainty. We expect decision-makers to be more influenced by applicants' social categories under these conditions.

Our key claim is that, given some uncertainty about who is the "best" candidate and a tendency to categorize applicants, decision-makers will give preferential treatment to candidates from the numerical majority. To place this claim in context of prior research, we will briefly review the judgment and decision making literature on the conditions under which people tend to account for sample sizes.

This literature shows that people as young as eleven (Piaget and Inhelder 1975) tend to have an intuitive understanding that, compared to small samples, large samples more closely approximate the distribution of an underlying population (e.g., Evans & Dusoir 1972; for a review, see Sedlmeier and Gigerenzer 1997). At first glance, this work appears to contradict findings showing that people do not adequately account for sample sizes in making judgments. For instance, a seminal paper by Kahneman and Tversky (1972) showed that people do not intuitively understand that distributional characteristics of samples, such as the mean value, exhibit larger variance in smaller samples. To explain these apparently contradictory findings, Sedlmeier and Gigerenzer (1997) emphasized the key distinction between two types of judgment tasks. Based on a review of 35 studies of intuitions about sample sizes, Sedlmeier and Gigerenzer concluded that people take sample size into account when a given large sample is compared to a given small sample (termed a "frequency distribution task" by Sedlmeier and Gigerenzer), but not when the distribution of a characteristic, such as the mean value, from a set of independent samples of a fixed size large is compared for a large vs. small sample size (a "sampling distribution task").

The situation in which we are interested - where a decision maker chooses a candidate from two give categories of applicants - is clearly of the former type, a "frequency distribution task." Thus, based on prior research, we assume that people will take sample sizes into account in such situations. Specifically, we assume that a decision maker will intuitively understand that the largest of two (different-sized) groups is more likely to contain the *best* candidate - as well as the *worst* candidate. This latter prediction – that decision makers will also assume that the majority category will also contain the worst candidate -- illustrates how our argument is distinct from a competing line of reasoning, namely that people will tend to view the majority candidates as better on average, or show a general preference for members of the majority.

It is important to note that it is not "irrational" for a decision-maker to expect that the best (or worst) candidate will be in the numerical majority category. Rather it is the tendency to categorize others based on characteristics (like race or gender) orthogonal to job performance *and then* to make inferences about candidate quality based on category membership size that results in an irrational outcome. If categories are orthogonal to job qualifications, a decision-maker would maximize his or her chance of finding the best candidate (and avoiding the worst one) by considering the *entire* pool of applicants, irrespective of category. But we will show that, even when the basis of categorization is unrelated to the task, decision makers nevertheless attend to categorical information and give preferential treatment to the majority category. Thus, by making use of irrelevant categorical information, they decrease the chances that they find the best one and, in the process, discriminate against numerical minorities.

Hypotheses and Studies

We test two hypotheses in two studies. First, we test our hypothesis that decision-makers attend to social categories. As a result, decision-makers *tend to expect a member of a larger category to be more likely to be the "best" candidate.* To distinguish our hypothesis from a tendency to view majorities as better overall, we also test the hypothesis that decision-makers will *expect a member of the numerical minority category to be less likely to be the worst candidate.* Our first test of these hypotheses is in a non-social domain (picking lottery tickets). The first study allows us to demonstrate that the hypothesized process is a general feature of how people think about categories. Study 2 tests the hypotheses in the context of ranking job applicants.

1. Study 1 Methods

1.1. Participants

A total of 960 users of Amazon's Mechanical Turk (https://www.mturk.com) took part in Study 1. Mechanical Turk is an online labor market that allows requesters to post tasks that users (or "workers") can complete for monetary rewards. Workers can be located anywhere, but the majority are from the US or India (Eriksson & Simpson 2010). As others have noted (Mason & Watts 2009; Paolacci, Chandler, & Ipeirotis 2010; Buhrmester et al. in press), Mechanical Turk provides a convenient and reliable data source for conducting certain types of behavioral studies.

In addition to the two hypothetical lottery tasks, respondents completed basic demographic information, including gender (45% female), age (median = 27), and level of education (37% bachelors degree).

1.2. Procedure and Materials

Participants were asked to imagine they were taking part in two hypothetical lotteries. The first lottery was composed of 19 buttons in one of two different colors: 15 in a majority color (e.g., blue) and 4 in a minority color (e.g., red). The instructions explained that a computer had randomly selected one winning button (worth $200) out of these nineteen. Each respondent was asked to imagine that he or she could select four buttons, and was asked what she thought would be a wise strategy "to maximize your chance of winning:" *to select most or all of the buttons from the majority color, to select most or all buttons from the minority color,* or "*it doesn't matter*." The correct, or normative, answer is that "it doesn't matter." That is, because the computer randomly selected one winning button, the color categories were irrelevant. At issue, then, is the extent to which people tend to misuse these categories. Specifically, we expect that many participants will make use of the categories and, among those who do, there will be a tendency to favor the majority color. That is, participants who make use of categorical information will tend to apply the law of large numbers to the two categories and thus expect that the larger category is more likely to contain the winning button.

After indicating what she thought would be the best strategy to maximize her choice of winning, we presented the participant with a second lottery. Here, participants were asked to imagine that the goal was to avoid picking a single *losing* button. (Participants were told that picking the losing button would result in a loss of $200.) Again, participants were asked to imagine that they would need to pick four buttons, and asked what they would consider the best strategy for avoiding a losing button: selecting most or all of the buttons from the majority color, selecting most or all of the buttons from the minority color, or that it doesn't matter which color buttons they selected. As for the first lottery, we emphasized that the losing button had been randomly selected via computer. Thus, the correct (normative) answer was that it doesn't matter which color one favored to avoid losing. Because our argument suggests that participants would use the (irrelevant) categorical information and tend to assume that the majority color would contain whatever it is they were looking for (in this case, the button they needed to avoid), we predicted that those who used categorical information would tend to select buttons from the *minority* color. Note that responses to the losing lottery allow us to distinguish our argument from an alternative prediction that participants simply tend to think the majority category is "better" or "luckier."

Finally, to ensure that the predicted lay statistical error was not restricted to those with no prior training in statistics, participants were asked to indicate their experience with statistics, including whether they had taken courses in statistics, and their self-reported ability with statistics and probabilities. These measures were taken on scales from 1 to 6. The entire procedure took about 10 minutes.

2. Study 1 Results

For the winning lottery question, less than half (42.5%) of participants responded correctly that "it doesn't matter" which buttons one picks. As predicted, the most common error was to favor the majority color. Of those who indicated that the best strategy would be to select buttons from a given category to maximize their chances of selecting the winning button, 65% indicated they would tend to select buttons from the majority color, compared to 35% who indicated they would favor the minority category, $\chi^2 = 48.73$, $p < .001$.[2]

[2] We assessed whether the tendency to make the predicted error (focusing on the majority color versus the minority color) depended on respondent's gender, age, level of education, or country of origin (US, India, or other). Only country of origin had a significant impact, such that US respondents were marginally more likely ($p = .06$, 2-tailed) than others to make the predicted error. 77.5 of US respondents said they would focus on the majority color vs.

The strategies for avoiding the losing button allow us to rule out the alternative explanation that participants simply considered the majority category better or luckier. Again, as for the winning lottery question, less than half (44.5%) of respondents responded correctly that it doesn't matter which buttons one chooses from. Our argument predicts that the remaining respondents will tend to assume that the majority category will be disproportionately likely to contain the losing button (in addition to the winning button). As expected, participants who gave incorrect answers were substantially more likely to favor the minority (58.4%) than majority (41.6%) category in order to avoid losing, $\chi^2 = 14.78$, $p < .001$. We did not find any effects of country of origin, gender, age, or education level on the tendency to pick the minority color to avoid losing (all *p*s > .4).

As additional evidence that responses to the "winning" and "avoid losing" questions stem from the same underlying process, there is substantial overlap between those participants who stated that they would select from the majority pile to win and those who stated they would select from the minority pile to avoid losing. For instance, among the 306 respondents who indicated they would favor the minority pile to avoid losing, more than 75% stated they would pick from the majority pile to win. Table 1 gives the crosstab for respondents' choices for the winning lottery and the losing lotteries. Other than the correct answer, the largest cell (containing 24.1% of responses) is the one corresponding to the lay-statistician choices (focusing on the majority category to win and the minority category to avoid losing).

{Table 1 about here.}

22.5% who said they would focus on the minority color ($p < .001$). Importantly, we still observe the predicted effect among the remaining (mostly Indian) respondents, among whom 59.3% focused on the majority color and 40.7% focused on the minority color ($p < .001$).

Importantly, those without prior training in statistics were not more likely to make the predicted lay-statistician error. Indeed, if anything, we observed a marginally significant tendency for those who reported having taken courses in statistics to attend to category size. Specifically, among respondents who either answered both questions correctly (by noting that it did not matter which buttons were selected to win and to avoid losing) or made the predicted lay error in both questions (by selecting the majority color to win and minority color to avoid losing)[3] we find that the proportion of lay errors was marginally lower among respondents *without* prior statistics training (34%), compared to those with some prior training (41%), $\chi^2 = 3.00$, p = .08.[4] Finally, those who answered both questions correctly did not differ from those who made the predicted lay errors either in self-reported ability in statistics (3.54 vs. 3.52, t (598) = .21, p = .83) or self-reported ability with probabilities (3.71 vs. 3.70, t (598) = .22, p = .83). These results suggest that our findings are not limited to people with less exposure to statistics.

3. Study 2 Methods

Given the strong support for our argument in Study 1, our second study tested the argument in a social domain. As noted earlier, we predict that participants will tend to view numerical majorities as more statistically representative, and therefore more likely to include both the best and worst candidates. Thus, Study 2 tests the hypothesis that persons charged with the task of evaluating job applications will tend to view *i*) applications from highly qualified numerical

[3] This excludes 37.7% and 36.6% of those with training and without training in statistics.

[4] Because this difference is only marginally significant and we did not predict it, we do not want to make too much of it. However, it is possible that those with training in statistics may be more apt to make the error because they are more apt to know that larger samples tend to be more representative of an underlying population.

majority applicants as better than equally qualified applications from numerical minority applicants, and *ii*) applications from relatively unqualified numerical majority applicants as worse than equally unqualified applications from numerical minority applicants.

3.1. *Participants*

A total of 641 Mechanical Turk users took part in Study 2. In addition to the main task described below, respondents completed demographic information about themselves. 27% of respondents were women, 55% had a bachelor's degree as their highest level of education, and median age was 27 years. Because none of these variables, nor country of origin, qualified our findings (all *p*s > .15), we do not discuss them any further.

3.2. Procedure and Materials

Respondents were asked to evaluate resumes for a Mechanical Engineering job vacancy, using employment advertisement materials developed in prior work (Foschi, Lai & Sigerson 1994). We sought to make the procedure as straightforward for participants as possible. Thus, rather than having each participant evaluate a large number of resumes, we pre-tested a pool of eighteen resumes. Ostensible applicants varied on educational qualifications (major, coursework, internships, GPA) and prior work experience.

Based on pre-tests, we selected the two applicants that were ranked as most qualified for the position and the two that were ranked as least qualified. We presented Study 2 participants with these four applications, along with category information about each applicant. Category information was presented as follows: “For statistical purposes, we created two categories of applications. These categories are based on demographic information that is not relevant for hiring decisions. Both categories were diverse with respect to other demographic characteristics, including gender and race/ethnicity. (We have removed this demographic information from applications, as it is not relevant to the job description.)" We used various (counter-balanced) labels for the majority- and minority- categories, e.g., Omega and Zeta; Red and Blue, etc. (See supplementary materials included for reviewers.)

The instructions then explained that we originally received 38 applications in the "majority" category and nine applications in the "minority" category, and that we pre-selected the best and worst applicant from each category. Respondents were asked to evaluate these four candidates against the job description and to rank-order them from most- to least-qualified. Applications included details of the job candidates' education, training, and prior work experience. (See supplementary materials included for reviewers.) We counter-balanced which applications were assigned to the majority versus minority categories, such that each application appeared equally often in each category.

Because we emphasized that the basis of categorization was irrelevant to the job description and participants were only asked to evaluate four applications, this task constitutes a relatively strong test of the hypothesis that participants will make use of the arbitrary categorical information.

4. Study 2 Results

Our analyses focus only on the 434 respondents who ranked all candidates, and who gave the objectively best (worst) candidates the highest (lowest) rankings. Of these respondents, 241 (55.6%) chose the top candidate from the majority and the same number chose the worst candidate from the majority. Each of these effects is statistically significant, $\chi^2 = 5.21$, $p < .05$. Thus, these results confirm our hypothesis: respondents tended to rank highly qualified majority candidates as better than equally qualified minority candidates, and relatively unqualified majority candidates as worse than equally qualified minority candidates.

Further evidence of the hypothesized impact of category size on participants' rankings of the applicants can be seen by comparing the responses of participants whose rankings of the best and worst candidates were from the same category. Among these participants, it was substantially more common for participants to think that both the best and worse candidate were in the majority, compared to the minority (60.4% vs. 39.6%, $\chi^2 = 10.02$, $p < .005$).

Summing up, the Study 2 findings support our argument linking lay statistical thinking to minority discrimination. We believe the Study 2 findings are especially noteworthy, given that the instructions explicitly emphasized the removal of information relevant to the categorization "as it is not relevant to the job description."

5. Discussion

As discussed above, traditional social science models have assumed that discrimination in labor market processes are driven by rational and explicit processes. For instance, statistical models of discrimination explain discrimination against African American candidates via employers' rational responses to the uncertainty inherent in hiring processes: given that it is often impossible to accurately predict how a prospective employee will perform, an employer may act on beliefs about the average talents and skills of African Africans vs. European Americans and hire what they believe to be "safer bet," i.e., white candidates.

These traditional approaches have shed much light on the processes that can generate discrimination in the labor market. But researchers have recently emphasized the need for alternative investigations into the processes through which discrimination occurs (NRC; Quillian cite). As Reskin (2002, 219) notes: “Existing theories of individual-level discrimination that emphasize people's conscious choices do not encompass common sources of micro acts of discrimination. We need to theorize discrimination more broadly to include automatic cognitive processes that distort information processing and decision making by individuals in ways that - unless checked-lead to micro (and sometimes macro) acts of discrimination.” To this end, recent research in social cognition and social psychology has yielded an array of new insights into the implicit and automatic processes governing discrimination. Among other things, this literature has provided powerful evidence of the automatic and pervasive tendency for employers to categorize applicants or candidates for promotion based on characteristics such as race and gender.

While this emerging literature on the implicit and automatic processes governing discrimination differs from traditional work on explicit and rational bases in a number of fundamental ways, the two streams of research both focus on how the *content* of categories impact discrimination. In contrast, the arguments and evidence introduced in this paper address how the relative *sizes* of the categories lead to discrimination against category members. That category size impacts discrimination is important because, as noted earlier, employers typically face situations where there are a large number of applications from majorities versus minorities, males versus females, etc.

Our basic argument is that, other things equal, employers will tend to assume that a social category that makes up a larger portion of the applicant pool (e.g., whites) will be most likely to contain what they are looking for (the best candidate). If salient categories are orthogonal to what the decision maker is searching for (talent, reliability, productivity), the lay-statistician mechanism introduced in this paper will tend to result in discrimination.

We tested and found support for our argument across two studies. To demonstrate that the proposed mechanism is a general feature of how people think about categories, we first tested our argument in a non-social domain. Study 1 showed a tendency for respondents to use categorical information (in this case, the color of lottery buttons) and to assume that the larger category was more likely to contain a single winning button *and* to contain a single losing button. Our second study tested the arguments in a labor market discrimination context. We found that respondents tended to rank highly qualified candidates from the numerical majority as better than otherwise identical candidates from the numerical minority and to rank relatively unqualified candidates from the numerical majority as worse than otherwise identical candidates from a numerical minority. Importantly, we observed these effects even though these categories had no prior meaning or history for respondents, and despite the fact that the instructions emphasized that the basis of categorization was irrelevant to the job.

Of course, this is only a first step and much remains to be done. Perhaps most importantly, we know from prior work that discrimination has multiple causes and we need to understand how the mechanisms introduced in this paper interact with mechanisms studied in prior work on the *content* of categories. For instance, under certain conditions, we might expect the mechanisms described here to combine with previously studied bases of discrimination to further disadvantage females and racial/ethnic minorities. This would be most likely when employers must evaluate an applicant pool containing a disproportionate share of whites or males, as typically happens for high ranking positions. Our argument implies that the employer would discriminate against numerical minorities (e.g., blacks) who may also be subject to discrimination based on mechanisms such as anti-black bias, racial animus, and statistical discrimination based on beliefs about the talents and skills of blacks versus whites.

On the other hand, all other things equal, when applicant pools are composed disproportionately of African Americans or females, our argument predicts discrimination patterns that can favor blacks and females. It is especially important for future research to identify how category size interacts with category content to impact discrimination, especially when the two mechanisms would, in isolation, generate contrasting effects.

We also need to better understand the conditions under which category size does not lead to discrimination. For instance, a substantial minority of participants in Studies 1 and 2 did not make use of categorical information. Future research should identify the personal, situational, and organizational contexts that decrease reliance on categorical information.

Although much remains to be done, we believe the arguments and findings introduced in this paper may be relevant to a range of issues. For instance, research by Holzer & Neumark (2000) shows that affirmative action policies increase the number of applications from females and minorities. Our argument thus suggests an additional mechanism through which these policies can reduce hiring discrimination: by reducing disparities in the relative *number* of applicants from majority vs. minority candidates, such policies should reduce discrimination based on the lay statistician mechanism outline above.

Related to the previous point, larger numbers of women in a given organization or occupation leads to further increases in the number of women in the occupation (Strober and Arnold 1987; Jolly et al. 1990; Cohen et al. 1998). Prior work has identified a number of reasons for this pattern (see Cohen et al. 1998). Our argument points to yet another mechanism through which this might occur: if women apply in larger numbers, they will make up a larger share of the pool of applicants for employment or promotion. Decision-makers may therefore disproportionately select females for positions.

Under the same logic, our argument may help explain the stubborn persistence of sex and race-segregation in some labor market domains. If women or blacks make up only a minority of a given industry, women and blacks will likely apply in fewer numbers (e.g., Cohen et al. 1998). If so, our argument suggests that, all other things equal, they will tend to be disproportionately excluded from jobs or promotion.

Like other category-based approaches to discrimination, our arguments suggest that hiring practices that mask the dimensions along which decision-makers typically categorize others (e.g., race or ethnicity and gender) may reduce discrimination. In contrast to prior work, we expect such policies would be most effective at combating discrimination when applicants are highly skewed (e.g., applicant pools with many males and few females, or many whites and few blacks).

In short, there are a number of issues that need to be addressed in future work. Yet we think the current research shows the value of addressing how cognitive size leads to discrimination. Broadening our understanding of the processes underlying discrimination should yield insight into the best practical solutions to problems of discrimination.

References

Aigner, D. J., & Cain, G. G. (1977). Statistical theories of discrimination in labor markets. *Industrial and Labor Relations Review, 30,* 175-87.

Bendick, M. Jr., Jackson, C., & Reinoso, V. (1994). Measuring employment discrimination through controlled experiments. *Review of Black Political Economy 23*, 25–48.

Bobo 2001 "Racial attitudes and recial relations at the close of the 20th century."

Browne, I., & Misra, J. (2003). The intersection of gender and race in the labor market. *Annual Review of Sociology, 29*, 487–513.

Buhrmester, M. D., Kwang, T., Gosling, S. D. (in press). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data?. *Perspectives on Psychological Science.*

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.

Dovidio & Gaertner 2004 A*versive Racism in Advances in Experimental social pscyhology*

Evans, J. St B.T. & Dusoir, A.E. (1977). Proportionality and sample size as factors in intuitive statistical judgment. *Acta Psychologica*, *41*, 129-137.

England, P. (1992). *Comparable worth: Theories and evidence*. Hawthorne, NY: Aldine de Gruyter.

Eriksson, K., & Simpson, B. (2010). Emotional reactions to losing explain gender differences in entering a risky lottery. *Judgment and Decision Making*, *5*, 159-163.

Roberto M. Fernandez and Mabel Botelho Abraham . 2010. "From Metaphors to Mechanisms: Gender Sorting In(to) an Organizational Hierarchy." MIT Sloan School Working Paper 4779-10

Fernandez, Roberto M., and Marie Louise Mors. 2008. "Competing for jobs: Labor queues and gender sorting in the hiring process." *Social Science Research* 37:1061-1080."

Firth, M. (1982.) Sex discrimination in job opportunities for women. *Sex Roles*, 8, 891-901.

Fiske, S. T. (1998.) Stereotyping, prejudice, and discrimination. In *Handbook of Social Psychology,* 4th ed., D. Gilbert, S.T. Fiske, and G. Lindzey, eds. New York: McGraw Hill.

Foschi, M., Lai, I. & Sigerson, K. (1994). Gender and double standards in the assessment of job applicants. *Social Psychology Quarterly*, *57*, 326-39.

Frieze, I. H., Olson, J. E., & Cain, D. (1990). Perceived and actual discrimination in the salaries of male and female managers. *Journal of Applied Social Psychology,* 20, 46–67.

Holzer, H. & Neumark, D. (2000). Assessing affirmative action. *Journal of Economic Literature, 38*, 483-568.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge University Press.

Macrae, C. N., & Bodenhausen, G. V. (2000). Social cognition: Thinking categorically about others. *Annual Review of Psychology, 51*, 93–120.

Mason, W., & Watts, D. (2009). Financial incentives and the "performance of crowds." *HCOMP '09: Proceedings of the ACM SIGKDD Workshop on Human Computation*, 77–85.

Pager, D., & Shepherd, H. (2008). The sociology of discrimination: Racial discrimination in employment, housing, credit, and consumer markets. *Annual Review of Sociology, 34*, 181–209.

Reskin, B. F. (1993). *Annual Review of Sociology …*

Reskin, B. F. (2003). Including mechanisms in our models of ascriptive inequality. *American Sociological Review*, *68*, 1-21.

Salancik, G.R. and Pfeffer, J. (1978.) Uncertainty, secrecy, and the choice of similar others. *Social Psychology* 41:246–55.

Schwab, S. (1986). Is statistical discrimination efficient? *American Economic Review, 76,* 228-234.

Stiglitz, J. E. (1973). Approaches to the economics of discrimination. *American Economic Review, 63*, 287-295

Tversky, A. & Kahneman, D. (1982). Judgments of and by representativeness. In D. Kahneman, P. Slovic & A. Tversky (Eds.), Judgment under uncertainty: Heuristics and biases. Cambridge, UK: Cambridge University Press.