Research Article

Psychological Science
2020, Vol. 31(8) 1013–1024

DOI: 10.1177/0956797620929979
www.psychologicalscience.org/PS
SAGE

On Intersectionality: How Complex Patterns of Discrimination Can Emerge From Simple Stereotypes

Neil Hester[1], Keith Payne[2], Jazmin Brown-Iannuzzi[3], and Kurt Gray[2]
[1]Department of Psychology, McGill University; [2]Department of Psychology and Neuroscience, University of North Carolina at Chapel Hill; and [3]Department of Psychology, University of Virginia

Abstract
Patterns of discrimination are often complex (i.e., multiplicative), with different identities combining to yield especially potent discrimination. For example, Black men are disproportionately stopped by police to a degree that cannot be explained by the simple (i.e., additive) effects of being Black and being male. Researchers often posit corresponding mental representations (e.g., intersectional stereotypes for Black men) to account for these complex outcomes. We suggest that complex discrimination can be explained by simple stereotypes combined with threshold models of behavior—for example, "if someone's threat level seems higher than *X*, stop that person." Simulations provide proof of this concept. We show how gender-by-race discrimination in both promotions and police stops can be explained by simple stereotypes. We also explore race-by-age discrimination in police stops, in which racial disparities are greater for young adolescents. This work suggests that complex behaviors can sometimes arise from relatively simple cognitions.

Keywords
intersectionality, theoretical models, threshold models, prejudice, decision rules, open data, open materials

Received 7/11/19; Revision accepted 4/9/20

Why are Black men disproportionately stopped by police? Why are White men three times more likely to become CEOs than Black women? Discrimination is often best described as the multiplicative combination of two or more identities. One intuitive explanation is that underlying cognitions closely match the form of outcomes. The term *intersectional stereotypes* refers to the overlap of multiple social identities that may combine in unique ways to contribute to discrimination (Crenshaw, 1989). The stereotype of criminality, for example, may be applied to Black men more strongly than can be explained by the additive effects of race and gender stereotypes (Thiem, Neel, Simpson, & Todd, 2019).

These intersectional stereotypes certainly exist (Ghavami & Peplau, 2013) and explain variance in discrimination (e.g., Hester & Gray, 2018; see Kang & Bodenhausen, 2015). However, positing a unique mental representation for each behavioral pattern is an unparsimonious approach to building theories. Psychological theory advances more quickly when a small number of unobserved entities can be used to explain a large range of behavioral outcomes (Epstein, 1984; Gawronski & Bodenhausen, 2015). We suggest that complex (i.e., multiplicative) patterns of discrimination can in principle result from simple (i.e., additive) stereotypes that combine with simple decision rules—for example, "if someone has a threat level of *X* or higher, that person will be stopped by police." This model offers a parsimonious potential explanation for complex patterns of behavior.

We draw on classic work describing how decision criteria and population distributions combine to explain behavioral outcomes. This class of models includes

Corresponding Author:
Neil Hester, McGill University, Department of Psychology, 2001 McGill College Ave., 14th Floor, Montreal, Quebec H3A 1G1, Canada
E-mail: neilrhester@gmail.com

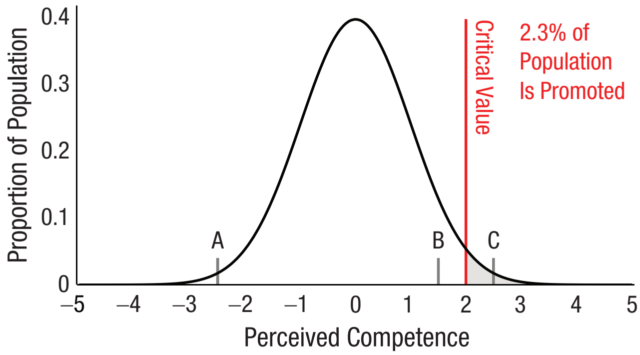

Fig. 1. A threshold model of an employer's decision to promote or not promote someone, assuming a normal distribution of applicants. The red line indicates the critical value of perceived competence at which a person is promoted, on average. In this case, the critical value means that 2.3% of a population is promoted. In this model, person B and person C are similar in perceived competence but experience different outcomes (person C is promoted, but person B is not); person A and person B are very different in perceived competence but experience the same outcome (neither is promoted). The normal distribution represents the distribution of perceived competence in the population. Solving for the area under the curve to the right of the threshold shows that 6.7% of the population above a critical value of 1.5 will be promoted.

signal detection theory (see Swets, Dawes, & Monahan, 2000), item response theory (see van der Linden & Hambleton, 2013), and threshold models (see Granovetter, 1978; Vallacher & Nowak, 1994). Threshold models form the basis of our proposed model.[1] Threshold models explain how, for example, two people with relatively similar levels of a trait (the perceived competence of person B and person C; see Fig. 1) can experience different outcomes (one is promoted, the other is not) and how two people with very different levels of the same trait (person A and person B) can experience the same outcome (not being promoted; see Fig. 1).

By comparing how the same critical value impacts outcomes for different populations (e.g., women vs. men), threshold models describe one way in which group differences in perceived traits (whether actual or stereotypical) translate to discrimination. Importantly, the pattern of discrimination that emerges from group-mean differences depends on the critical value. In Figure 2, using the green critical value results in 0.85 women being promoted per man (e.g., to a manager-level position), whereas using the red critical value results in 0.62 women being promoted per man (e.g., to an executive-level position). Thus, the same gender stereotype for perceived competence can result in different disparities depending on the critical value (which itself is stereotype independent).

The relative position of the critical value can also shift as the function of a second population effect. Imagine comparing manager promotion rates (green line in Fig. 3) for men and women who have either 1 or 10 years of experience. This results in 0.85 women being promoted per man at 10 years of experience, but only 0.62 women being promoted per man at 1 year of experience—the same ratios as in the previous example.

In this research, we leveraged threshold models to suggest that complex patterns of discrimination can result from simple stereotypes. We first explain how multiple-population threshold models can explain discrimination. Then we present the threshold models we used to test whether three complex patterns of discrimination might be accounted for by additive stereotypes. Our hope is that this demonstration improves theorizing by reducing the need to posit a unique mental representation for every complex pattern of discrimination.

Threshold Models

Basic threshold models

Decisions about how to treat other people are often binary—a police officer cannot half-arrest someone. As with continuous decisions, people make binary decisions on the basis of their perceptions of diagnostic traits (e.g., threat for police stops). However, unlike in continuous decisions, the link between perceived traits and binary outcomes is nonlinear (as Fig. 1 demonstrates). Threshold models of behavior (Cox, 1987; Granovetter, 1978; Vallacher & Nowak, 1994, 1997) describe how perceptions of diagnostic traits predict binary outcomes. Imagine that an employer is deciding which workers to promote (Y) based on whether their

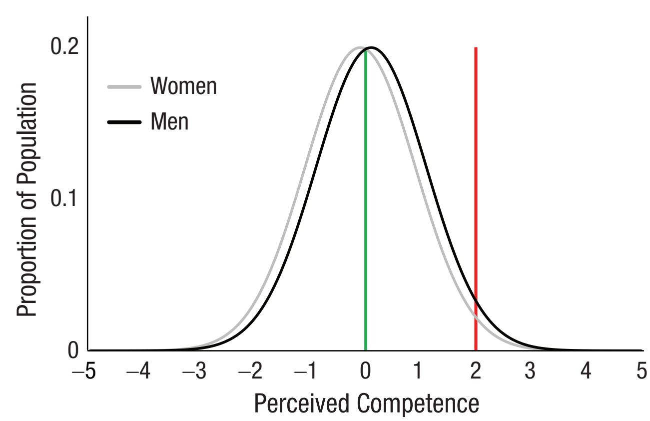

Fig. 2. A threshold model with population distributions for both women (M = .1, SD = 1) and men (M = .1, SD = 1). Two critical values for promotion result in two different patterns of gender disparity. For the central critical value of 0 (green), 0.85 women are promoted per man, and 7.97 fewer women are promoted than men (per 200 employees). For the more extreme critical value of 2 (red), 0.62 women are promoted per man and 1.10 fewer women are promoted than men (per 200 employees).

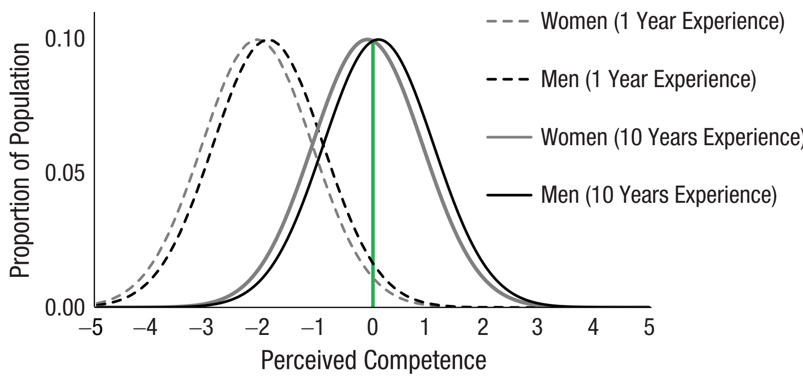

Fig. 3. A threshold model with population distributions at 10 years of experience for women ($M = -.1$, $SD = 1$) and men ($M = .1$, $SD = 1$) and at 1 year of experience for women ($M = -2.1$, $SD = 1$) and men ($M = -1.9$, $SD = 1$). The green line shows the central critical value. In the 10-year group, 0.85 women are promoted per man, and 7.97 fewer women are promoted than men (per 200 employees). In the 1-year group, 0.62 women are promoted per man, and 1.10 fewer women are promoted than men (per 200 employees).

perceived competence (X) exceeds a critical value (c). Workers with more competence than c are promoted ($Y = 1$), whereas workers with less competence than c are not promoted ($Y = 0$):

$$Y = \begin{cases} 0 \text{ for } X < c \\ 1 \text{ for } X \geq c \end{cases}$$

The population-level consequences of a threshold model can be calculated by solving for the distribution's integral. Assuming a normal distribution, one can calculate the percentage of people who experience an outcome:

$$D(Y) = \int_c^{\infty} X \sim N(\mu, \sigma^2)$$

Four factors determine the percentage of people in a population (Y) who experience a binary outcome based on a perceived trait (X): mean (μ), standard deviation (σ), distribution shape, and critical value (c).

Inequity threshold models

Threshold models that include multiple population distributions can reveal how population differences in perceived diagnostic traits (whether driven by stereotypes or actual differences) translate to disparities in binary outcomes. Although group means differ in inequity threshold models, the same critical value is used to calculate the percentage of people in each group who experience an outcome. These percentage values translate to *relative disparity* between groups (e.g., "How many women are promoted per man?"), which corresponds to the percentages often used to quantify discrimination (e.g., "95% of CEOs are men"). The relative disparity between groups j and k is defined as the ratio of two threshold models:

$$\int_c^{\infty} X \sim N(\mu_j, \sigma_j^2) : \int_c^{\infty} X \sim N(\mu_k, \sigma_k^2)$$

The *absolute disparity* between two groups (e.g., "per n people, how many fewer women than men are promoted?") can also be derived. Absolute disparities represent the total impact of a group difference (e.g., "There are 827 more male CEOs than female CEOs"). The absolute disparity is calculated by multiplying each threshold model by group population (n) and taking the difference:

$$(n_j)\int_c^{\infty} X \sim N(\mu_j, \sigma_j^2) - (n_k)\int_c^{\infty} X \sim N(\mu_k, \sigma_k^2)$$

The relation between critical values and relative versus absolute disparities

Imagine a society in which women are stereotyped as 0.2 units less competent than men. Different critical values correspond to different promotion levels: one that requires moderate competence (e.g., manager; green line in Fig. 2), whereas the other requires outstanding competence (e.g., executive; red line in Fig. 2).

For managers, the relative disparity is small (0.85 women promoted per man). However, because so many people in the population exceed this critical value, the absolute disparity is large (7.97 more men

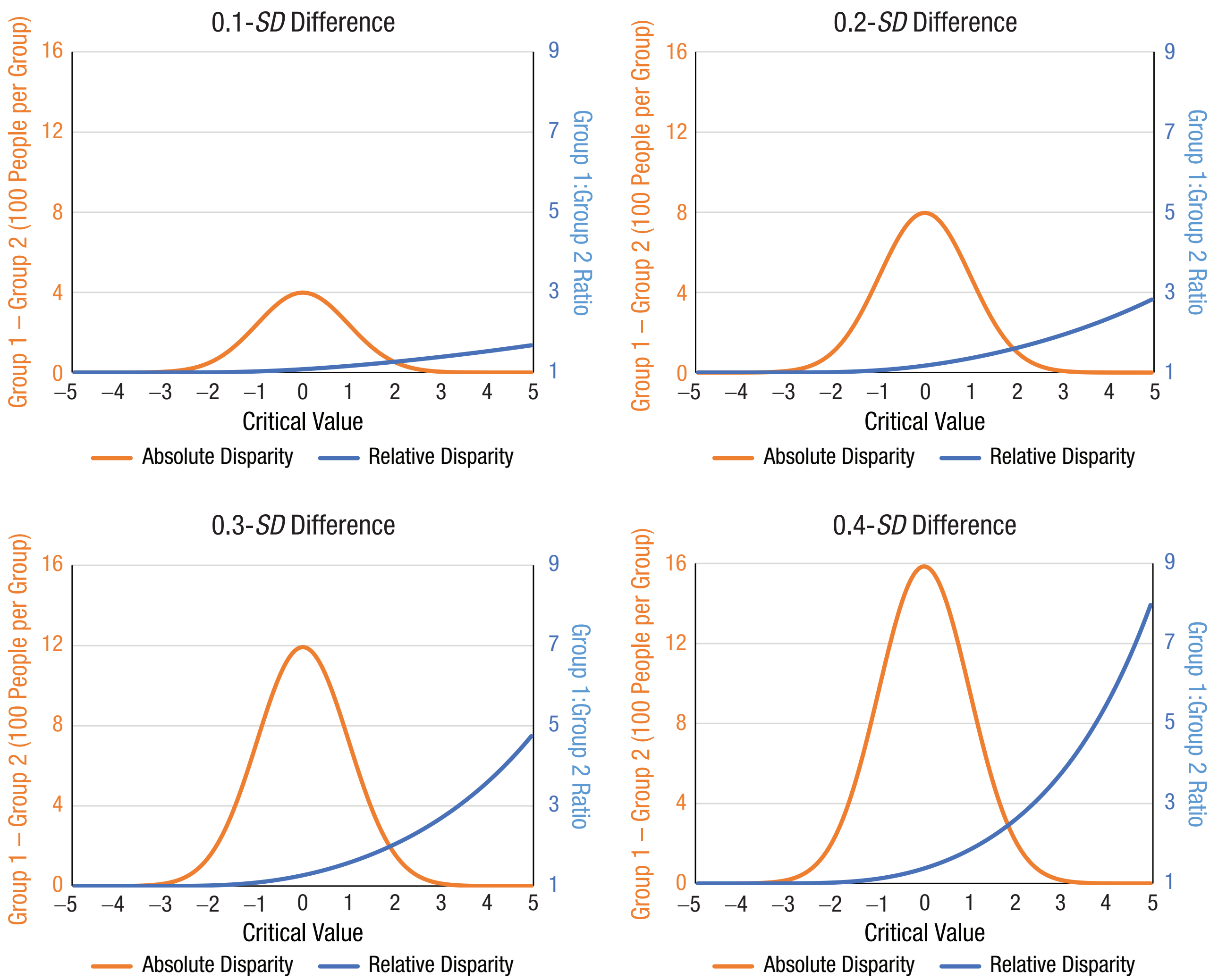

Fig. 4. Absolute and relative disparities graphed as a function of the position of the critical value, assuming that the groups are equal in size, the absolute mean across groups is 0, the standard deviation is 1, and the distribution is normal. Disparities are graphed for four magnitudes of group stereotyping: 0.1 *SD*, 0.2 *SD*, 0.3 *SD*, and 0.4 *SD*. As the magnitude of group stereotyping increases, absolute disparities increase in an approximately linear fashion (i.e., the apex of the normal distribution becomes higher), whereas relative disparities increase in an exponential fashion.

promoted than women per 200 employees). Conversely, for executives, the absolute disparity is small (1.10 more men promoted than women per 200 employees), but the relative disparity is large (0.62 women promoted per man). More generally, as the critical value moves right from the center, absolute disparities decrease and relative disparities increase (see Fig. 4). This distinction clarifies how disparities can simultaneously be smaller and larger for one outcome compared with another.

The Present Research

We explored patterns of complex discrimination often explained through multiplicative (i.e., intersectional) stereotypes. Using statistical simulations, we examined whether these complex outcomes can be explained by simple (i.e., additive) stereotypes combined with threshold models. Importantly, we sought to show only a plausible proof of concept that complex patterns of discrimination can emerge from simple stereotypes. We first examined the overrepresentation of White men as executives and then examined the overrepresentation of Black men as targets of stop-and-frisk policies. In both cases, we found evidence that complex patterns of race and gender can be explained by simple stereotypes combined with threshold models. Finally, we extended these models to the combination of race and age, exploring why relative racial disparities in police stops are higher among 12-year-old adolescents than among 20-year-old adults. Though this pattern appears to suggest intersectional stereotyping, we found that additive stereotypes combined with threshold models naturally explained this complex pattern of discrimination.

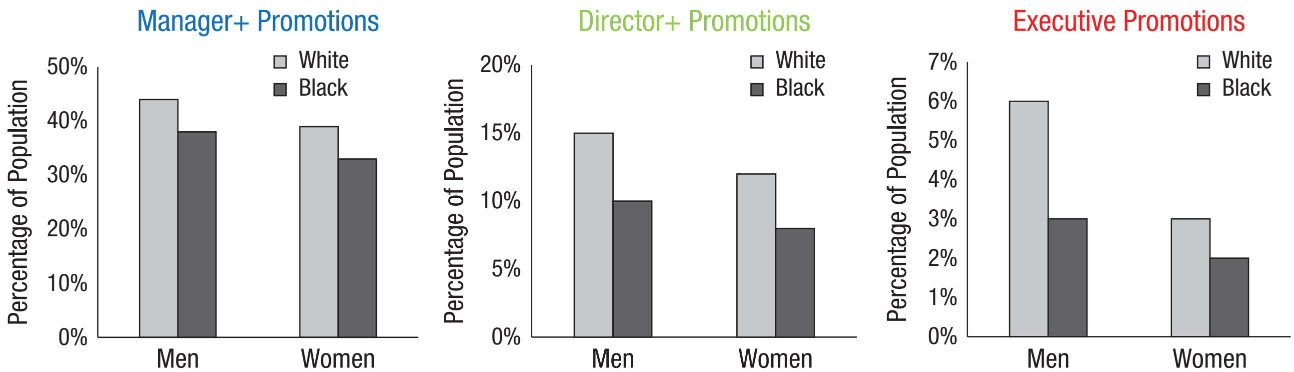

Fig. 5. Actual percentage of Black and White men and women promoted to manager or higher, director or higher, and executive (Simulation 1).

Modeling Approach

To test whether complex patterns of discrimination can be statistically explained through simple group stereotypes and threshold models, we used an optimization procedure to find the combination of group differences and critical values that would produce the pattern of outcomes that most closely matches the observed pattern. Then we calculated the difference between observed and predicted outcomes to evaluate the fit between the model and data using a chi-square test. If the fit between the model and data is close enough that they are not significantly different from each other, then the observed data are adequately explained by the defined model (e.g., Barrett, 2007). This approach to evaluating the fit between the model and the data yields an additional benefit: It allows one to estimate the magnitude of stereotyping that would result in the observed patterns of discrimination. All estimated solutions are provided at the Open Science Framework (https://osf.io/kcn63/).

Simulation 1: Gender Disparities Along the Corporate Ladder

Intersectional disparities in the workplace have received increasing attention from researchers and the popular press (Paul, Hamilton, & Darity, 2018; Tugend, 2018). Black women are uniquely disadvantaged—and White men uniquely advantaged—in part because of intersectional stereotypes (Maume, 2004; Rosette, Koval, Ma, & Livingston, 2016). We consider whether simple race and gender stereotypes about competence can be combined with threshold models to mirror observed race-by-gender patterns of discrimination in corporate promotions.

Data

For our observed outcomes, we used a 2019 survey of over 1,800,000 corporate employees conducted by the research team of PayScale (2019; see also Gruver, 2019). This survey provides data on the percentage of Black and White men and women promoted to manager, director, and executive. The large sample ensured that percentage estimates were accurate.

Among White men, 44% received promotion to manager or higher, 15% received promotion to director or higher, and 6% received promotion to executive or higher. Among White women, 38% received promotion to manager or higher, 10% received promotion to director or higher, and 3% received promotion to executive or higher. Among Black men, 40% received promotion to manager or higher, 12% received promotion to director or higher, and 3% received promotion to executive or higher. Among Black women, 33% received promotion to manager or higher, 8% received promotion to director or higher, and 2% received promotion to executive or higher (see Fig. 5).

Estimated solution

We hypothesized that the observed pattern of outcomes could be predicted by five variables: an overall gender difference for perceived competence, an overall race difference for perceived competence, a critical value for promotion to manager, a critical value for promotion to director, and a critical value for promotion to executive.

We used an optimization procedure to find the combination of these five values that best matched the observed outcomes. We predicted that the chi-square value for the model would support the null hypothesis that the predicted outcomes are not different than the observed outcomes. To scale the variables, we held standard deviations constant at 1 (i.e., groups had equal standard deviations) and set the intercept for women (i.e., the mean of perceived competence for women) to 0. Note that the solution is not sensitive to where the intercept is set; one could also fix the critical value and estimate the intercept instead.

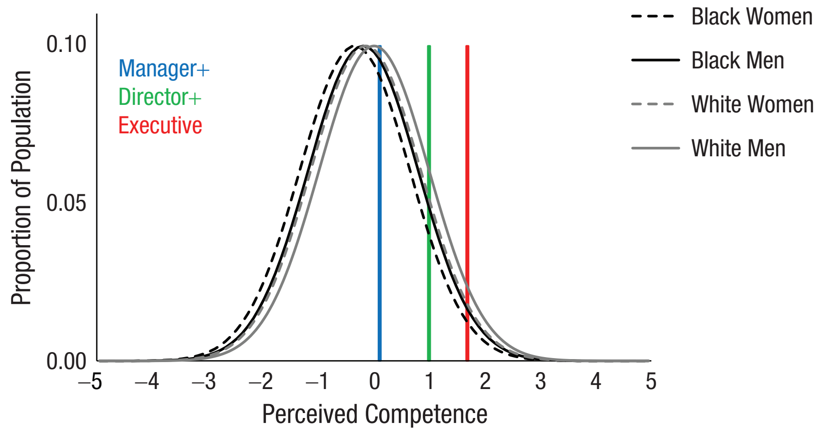

Fig. 6. Estimated solution for Simulation 1. The model estimated values for the gender difference in perceived competence, the racial difference in perceived competence, and the critical values for being promoted to manager or higher, director or higher, and executive.

The best-fitting solution estimated the overall gender difference in perceived competence at 0.21 standard-deviation units and the overall race difference in perceived competence at 0.14 standard-deviation units. Combined with three critical values at 0.11 (manager), 1.05 (director), and 1.65 (executive), this gender difference in perceived competence predicted a pattern of binary outcomes nearly identical to the observed pattern, $\chi^2(7) < 0.001$, $p > .99$ (Fig. 6). This solution demonstrates that race-by-gender discrimination in promotions does not require unique race-by-gender stereotypes for explanation.

Simulation 2: Race-by-Gender Disparities in New York Police Department (NYPD) Stops

Black males are disproportionately mistreated by police (Prison Policy Initiative, 2019), leading some researchers to suggest a unique intersectional threat-related stereotype for Black men. Although such a stereotype is plausible, here we tested whether simple threat stereotypes for race and gender can predict these complex patterns of discrimination in NYPD police stops.

Data

For our observed outcomes, we drew on 7 years of NYPD stop-and-frisk data—2006 to 2012—to examine race-by-gender stop patterns (data are available at www1.nyc.gov/site/nypd/stats/reports-analysis/stopfrisk.page). Given that previous work on racial stereotyping and discrimination has largely focused on non-Hispanic populations, we included only non-Hispanic Black and White stops, resulting in a sample of 2,418,931 stops. The large sample ensured that stop percentages were accurate representations of systemic patterns (i.e., they represent the behaviors of hundreds of different police officers).

In our data, Black men constituted 78.4% (N = 1,896,435) of the stops, White men constituted 14.2% of the stops (N = 343,125), Black women constituted 5.8% of the stops (N = 139,870), and White women constituted 1.6% of the stops (N = 39,501; see Fig. 7). Notably, the magnitude of discrimination in this data is especially pronounced given the demographics of New York City. When one considers only Black and White non-Hispanics, 36% of the population is Black, and 64% of the population is White (U.S. Census Bureau, 2010).

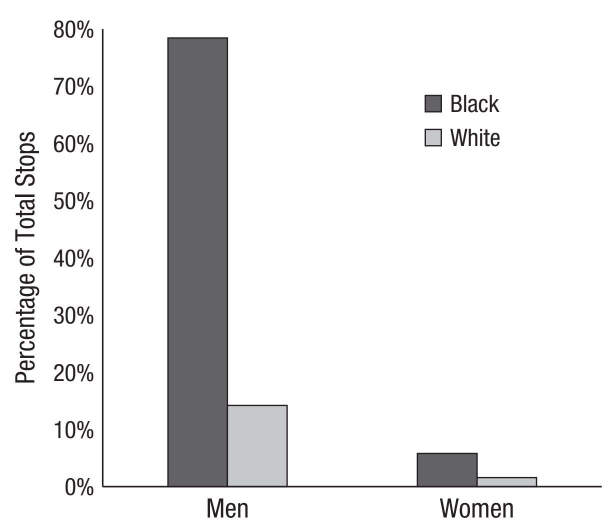

Fig. 7. Percentage of total stops by the New York City police under the stop-and-frisk program, 2006 to 2012, separately for Black and White men and women (Simulation 2).

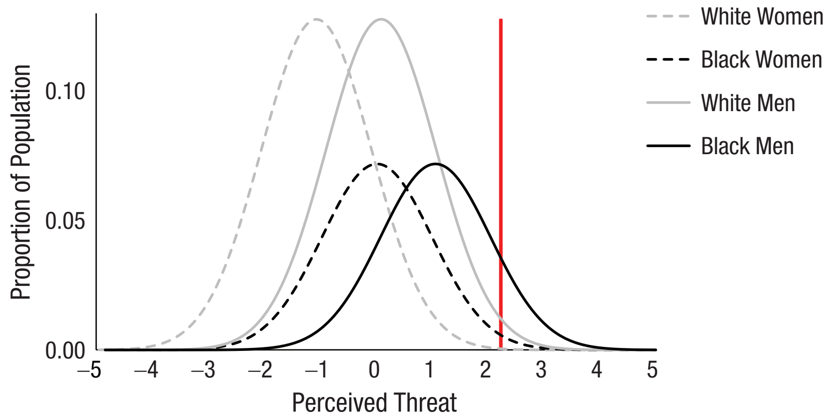

Fig. 8. Estimated solution for Simulation 2. The model estimated values for the racial difference in perceived threat, the gender difference in perceived threat, and the critical value for being stopped by police (red line).

When estimating our solution, we factored these demographics into the model.

Estimated solution

We hypothesized that the observed pattern of outcomes can be accounted for by the following variables: a main effect of race, a main effect of gender, and a critical value for the level of perceived threat at which the average police officer stops an individual.

We used an optimization procedure to find the combination of these three values that best matched the observed outcomes. We predicted that the chi-square value for the model would support the null hypothesis that the predicted outcomes are not different than the observed outcomes. To scale the variables, we held standard deviations constant at 1 (i.e., groups had equal standard deviations) and set the intercept for Black men (i.e., the mean of perceived competence for women) to 1. Again, the solution is not sensitive to where the intercept is set; one could fix the critical value and estimate the intercept instead.

Finally, a limitation of the stop-and-frisk data set—not including cases for individuals who were not stopped—required us to set a start value for the percentage of Black men in the population who were stopped by police. For our initial solution, we set this value at 10%. However, to ensure that our findings were not the result of researcher degrees of freedom, we reported robustness checks after the primary analysis to generalize our findings to an entire range of percentage values.

The best-fitting solution estimated the overall race difference in perceived threat at 0.99 standard-deviation units and the overall gender difference in perceived threat at 1.06, with a critical value of 2.30. These simple differences in perceived threat predicted a pattern of binary outcomes that closely matched the observed pattern, $\chi^2(1) < 0.83$, $p = .36$ (Fig. 8). This solution demonstrates that race-by-gender discrimination in police stops can be explained by simple race and gender stereotypes.

Robustness check

The one major researcher degree of freedom in this simulation was choosing a start value for the percentage of Black men in the population who are stopped (s). To ensure that our model did not predict the observed values only at specific values of s, we estimated solutions at various values of s. The predicted values adequately matched the observed values (i.e., $p > .05$ in chi-square tests) for values of s less than 27.7%, suggesting that main effects of race and age can account for the observed stop values under a wide range of assumed start values, alleviating concerns about researcher-set parameters. One additional consequence of this robustness check is that it shows how these models can fail to reproduce patterns using only main effects; that is, patterns of binary outcomes can occur that cannot be adequately accounted for by inequity threshold models.

Simulation 3: Race-by-Age Disparities in NYPD Police Stops

The data from NYPD's stop-and-frisk program show a complex race-by-age pattern, in which relative racial disparities are higher for young adolescents (e.g., 12–14 years old) than they are for young adults (e.g., 18–20

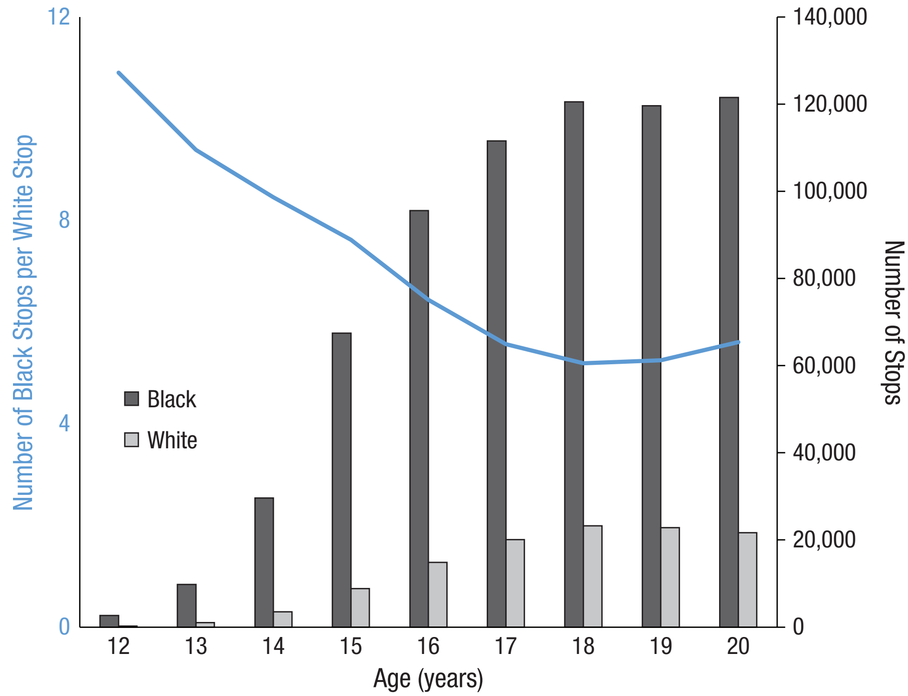

Fig. 9. Observed ratio of Black stops to White stops (left *y*-axis) alongside the observed number of Black and White stops (right *y*-axis) for individuals between the ages of 12 and 20. Observations are drawn from New York Police Department stop-and-frisk records.

years old). We tested whether this pattern of results can be explained using threshold models rather than intersectional stereotypes.

Data

For our observed outcomes, we drew on 7 years of NYPD stop-and-frisk data—2006 to 2012—to examine race-by-age stop patterns (data available at www1.nyc.gov/site/nypd/stats/reports-analysis/stopfrisk.page). In Figure 9, we provide observed data for stopped individuals ages 12 to 20. Given that previous work on racial stereotyping and discrimination has largely focused on non-Hispanic populations, we included only non-Hispanic Black and White stops, resulting in a sample of 794,704 stops. The large sample ensured that stop percentages were accurate representations of systemic patterns (i.e., they represent the behaviors of hundreds of different police officers).

We chose to focus on ages 12 to 20 for three reasons. First, we were concerned with the transition between adolescence and adulthood; developmental work roughly defines adolescence as the beginning of puberty (mean age in North America = 11.8 years; Westwood & Pinzon, 2008) to ages 18 to 20, at which point youths transition to young or emerging adulthood (Arnett, 2007). Second, there were too few stops for individuals 11 years old and under to accurately model. Third, the number of stops by age peaked at ages 18 to 20 and then began to decline. One explanation for this is that so many Black men are incarcerated in this period of their lives that the population of free Black men noticeably shrinks (one in five people incarcerated for 10 or more years is a Black man younger than 25; Urban Institute, 2017).

In the observed data, the relative disparity decreased as age increased. Police stopped 5.6 Black 20-year-olds for every 1 White 20-year-old, in line with previously documented racial disparities. However, relative racial disparities were much greater for adolescents: Police stopped 10.9 Black 12-year-olds for every 1 White 12-year-old.

However, absolute racial disparity is greater for 20-year-olds, because more of them are stopped overall. Police stopped 99,882 more Black 20-year-olds than White 20-year-olds, but stopped only 2,417 more Black 12-year-olds than White 12-year-olds. This pattern of data reflects a complex race-by-age interaction predicting police stops.

Notably, the magnitude of discrimination in this data is especially pronounced given the demographics of New York City. When one considers only Black and

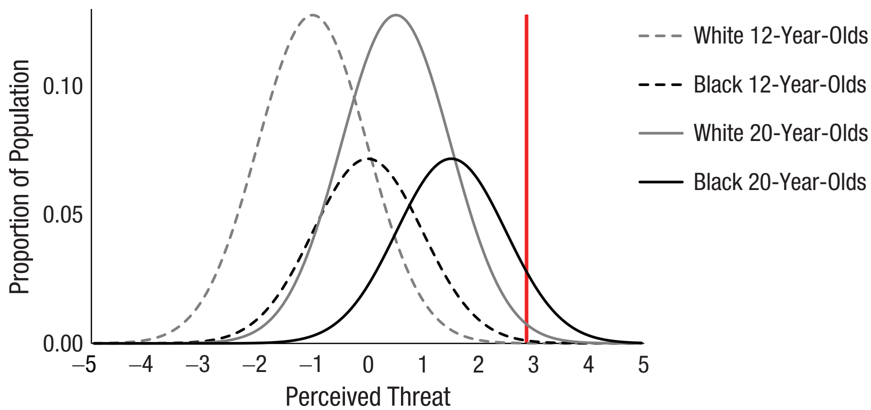

Fig. 10. Estimated solution for Simulation 3. The model estimated values for the racial difference in perceived threat, the age difference in perceived threat, and the critical value for being stopped by police (red line).

White non-Hispanics, 36% of the population is Black, and 64% of the population is White (U.S. Census Bureau, 2010). When estimating our solution, we factored these demographics into the model.

Estimated solution

We hypothesized that a main effect of race and a main effect of age on perceived threat could explain the observed complex pattern of police stops. We hypothesized that the observed pattern of outcomes could be predicted by the following values: a main effect of race on perceived threat, main effects of age on perceived threat (increasingly large effects as age increases, with separate age effects estimated for ages 12 through 20), and a single critical value at which someone is threatening enough to be stopped by the average police officer. We used an optimization procedure to find the combination of these values that best matched the observed outcomes, predicting that the resulting chi-square value for this predictive model would support the null hypothesis that the predicted outcomes are the same as the observed outcomes. The standard deviation was set to 1 for all groups, and the intercept (the mean threat for 12-year-old Black boys) was set to 0.

Finally, a limitation of the stop-and-frisk data set—not including cases for individuals who were not stopped—required us to set a value for the percentage of Black 12-year-olds in the population who were stopped by police. The data indicate that 45.7 times as many Black 20-year-olds were stopped as Black 12-year-olds. For the reported model, we assumed that 0.2% of Black 12-year-olds were stopped by police, which implies that 9.1% of Black 20-year-olds (0.2 multiplied by 45.7) were stopped by police. Values higher than 1% for the percentage of Black 12-year-olds stopped are infeasible (e.g., 2% of Black 12-year-olds being stopped would imply that 92% of Black 20-year-olds are stopped). Conversely, very low values, such as 0.01%, would suggest that only 0.46% of Black 20-year-olds are stopped, which is unlikely given reports on stop-and-frisk. For example, in a Pew Research Center (2016) poll, 18% of Black people reported being unfairly stopped by police in the last 12 months. To ensure that our findings were not the result of researcher degrees of freedom, we also conducted robustness checks after the primary analysis to generalize our findings to an entire range of percentage values.

The model estimated the overall race difference in perceived threat at 1.01 points and the maximum age difference in perceived threat (i.e., the difference between 12-year-olds and 20-year-olds) at 1.52 points, with a critical value of 2.86. Because the standard deviation was held constant at 1, the race and age differences in perceived competence can be reasonably interpreted as 1.33 standard deviations and 2.04 standard deviations, respectively. The distributions for Black and White 12-year-olds and 20-year-olds are illustrated in Figure 10.

The full solution included an effect estimate for every single age group (i.e., ages 12 through 20). This full solution produced predicted values that closely matched the observed values, $\chi^2(8) = 0.90$, $p > .95$. The estimated effects of age are provided in Figure 11 using 12-year-olds as the reference group. This curvilinear pattern of age effects is consistent with development research on puberty and adolescence: Physical changes associated with puberty are also associated with perceived threat (e.g., Wilson, Hugenberg, & Rule, 2017) and primarily shift between ages 12 and 16 (Westwood & Pinzon,

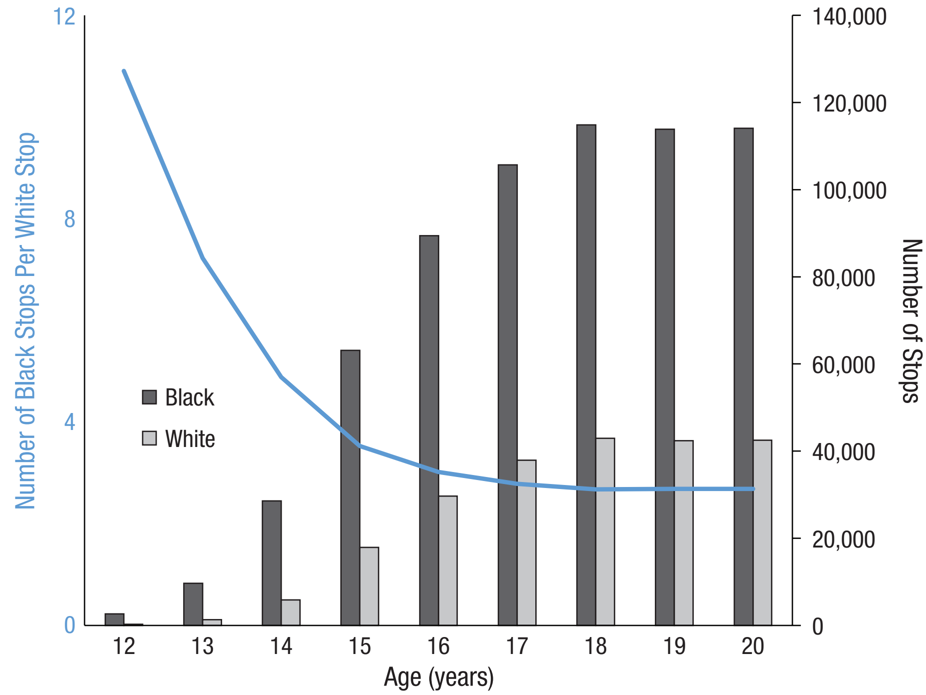

Fig. 11. Estimated ratio of Black stops to White stops (left *y*-axis) alongside the estimated number of Black and White stops (right *y*-axis) for individuals between the ages of 12 and 20. These values were derived from the estimated solution in Simulation 3 by scaling the estimated values such that the absolute number of Black 12-year-olds stopped matched the observed data.

2008). Overall, this solution suggests that the observed patterns of race-by-age discrimination in police stops do not necessitate race-by-age intersectional stereotypes. In Figure 11, we graph both the relative and absolute disparities predicted by our model, to allow for visual comparison with the observed values.

Robustness check

Again, a researcher degree of freedom in this study was choosing a start value for the percentage of Black 12-year-olds in the population that were stopped (*s*). To ensure that our model did not predict the observed values only at specific values of *s*, we estimated solutions at various values of *s*. The predicted values adequately match the observed values (i.e., $p > .05$ in chi-square tests) for values of *s* less than 1.35%. This suggests that simple effects of race and age can account for the observed stop values under a wide range of start values, alleviating concerns about researcher-set parameters.

General Discussion

Our simulations demonstrated that complex patterns of discrimination for binary outcomes can emerge from simple group stereotypes combined with threshold models. More broadly, the work shows that complex patterns of behaviors do not necessarily constitute evidence for equally complex cognitions. This insight may help psychologists develop accurate and parsimonious theories of stereotyping and discrimination, especially given the field's increasing focus on understanding diverse, intersecting identities (Cole, 2009; Goff & Kahn, 2013; Kang & Bodenhausen, 2015).

We also distinguished two complementary types of disparity—relative and absolute—and described how both result from underlying stereotypes. Researchers and policymakers often focus on relative disparities (e.g., "six Black men stopped per White man") because they intuitively capture inequality. However, in some cases, absolute disparities might be more practically useful. For example, a training intervention to reduce police officers' racial bias toward young adults—rather than young adolescents—would impact a higher number of people, despite larger relative racial disparity for young adolescents. Because relative and absolute disparities have different implications—both for practical interventions and theoretical development—researchers should clarify what kind of discrimination they are investigating or discussing.

In our models, we made certain statistical assumptions to argue our theoretical point. One such assumption is that populations follow normal distributions on perceived traits. Rather than view this assumption as a limitation, we view this choice as a necessary constraint because tests of fit require fixed parameters to allow degrees of freedom. One exciting implication of threshold models is that differences in distribution shape (skewness, kurtosis, general form) can produce group differences in the outcome independent of the mean (e.g., Hyde & Mertz, 2009; Strand, Deary, & Smith, 2006).

Multiple models for explaining complex discrimination

Although we focused on how relatively simple stereotypes can explain complex patterns of discrimination, we nevertheless stress that one theory or model need not account for all the variance in behaviors. Our theory does not oppose or refute theories of intersectionality; instead, we believe that the cognitions underlying patterns of discrimination likely include both additive stereotypes (which translate to behaviors via decision thresholds) and intersectional stereotypes (which can account for variance in behavior that cannot be fully explained by threshold models). The results of Study 3 illustrate this point: The estimated solution fitted the observed data closely but did not fully account for the observed data. This unexplained variance in the observed data may well be the result of intersectional stereotyping. We hope that theories of discrimination might incorporate both inequity threshold models and intersectional stereotyping to create a richer, more accurate account of when and why discrimination occurs.

Conclusion

Intersectionality has deeply contributed to our understanding of discrimination, helping us understand how and why certain groups are disproportionately mistreated. However, intersectional discrimination need not arise from intersectional stereotypes. In some cases, simple stereotypes can give rise to nuanced patterns of discrimination that appear more cognitively complex than they actually are. We hope this insight will guide future work on how stereotypes influence who is hired, fired, and stopped by police, as well as other outcomes that shape people's lives.

Transparency

Action Editor: Wendy Berry Mendes
Editor: D. Stephen Lindsay
Author Contributions

N. Hester developed the study concept, and K. Payne and K. Gray helped refine it. N. Hester performed the statistical modeling with assistance from K. Payne. N. Hester drafted the manuscript, and K. Payne, J. Brown-Iannuzzi, and K. Gray provided critical revisions. All authors approved the final version of the manuscript for submission.

Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Funding

This research was supported by the Russell Sage Foundation (Award No. 93-16-08).

Open Practices

Data for Simulation 1 were obtained from https://archive.vn/e5vCe and https://www.payscale.com/data/racial-wage-gap-for-men and are archived at https://osf.io/kcn63/. Data for Simulations 2 and 3 were obtained from www1.nyc.gov/site/nypd/stats/reports-analysis/stopfrisk.page. All models have been made publicly available via the Open Science Framework and can be accessed at https://osf.io/kcn63/. The design and analysis plans for the simulations were not preregistered. The complete Open Practices Disclosure for this article can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797620929979. This article has received the badges for Open Data and Open Materials. More information about the Open Practices badges can be found at http://www.psychologicalscience.org/publications/badges.

ORCID iD

Neil Hester https://orcid.org/0000-0003-3025-766X

Note

1. See this article's Open Science Framework project (https://osf.io/kcn63/) for an explanation of how threshold models differ from signal detection and item response models.

References

Arnett, J. J. (2007). Emerging adulthood: What is it, and what is it good for? *Child Development Perspectives*, *1*(2), 68–73. doi:10.1111/j.1750-8606.2007.00016.x

Barrett, P. (2007). Structural equation modelling: Adjudging model fit. *Personality and Individual Differences*, *42*, 815–824. doi:10.1016/j.paid.2006.09.018

Cole, E. R. (2009). Intersectionality and research in psychology. *American Psychologist*, *64*, 170–180. doi:10.1037/a0014564

Cox, C. (1987). Threshold dose-response models in toxicology. *Biometrics*, *43*, 511–523. doi:10.2307/2531991

Crenshaw, K. (1989). Demarginalizing the intersection of race and sex: A Black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. *University of Chicago Legal Forum*, *1989*, 139–168.

Epstein, R. (1984). The principle of parsimony and some applications in psychology. *The Journal of Mind and Behavior*, *5*, 119–130.

Gawronski, B., & Bodenhausen, G. V. (2015). *Theory and explanation in social psychology*. New York, NY: Guilford Press.

Ghavami, N., & Peplau, L. A. (2013). An intersectional analysis of gender and ethnic stereotypes: Testing three hypotheses. *Psychology of Women Quarterly*, *37*, 113–127. doi:10.1177/0361684312464203

Goff, P. A., & Kahn, K. B. (2013). How psychological science impedes intersectional thinking. *Du Bois Review: Social Science Research on Race*, *10*, 365–384.

Granovetter, M. (1978). Threshold models of collective behavior. *American Journal of Sociology*, *83*, 1420–1443. doi:10.1086/226707

Gruver, J. (2019, May 7). Racial wage gap for men. *PayScale*. Retrieved from https://www.payscale.com/data/racial-wage-gap-for-men

Hester, N., & Gray, K. (2018). For Black men, being tall increases threat stereotyping and police stops. *Proceedings of the National Academy of Sciences, USA*, *115*, 2711–2715. doi:10.1073/pnas.1714454115

Hyde, J. S., & Mertz, J. E. (2009). Gender, culture, and mathematics performance. *Proceedings of the National Academy of Sciences, USA*, *106*, 8801–8807. doi:10.1073/pnas.0901265106

Kang, S. K., & Bodenhausen, G. V. (2015). Multiple identities in social perception and interaction: Challenges and opportunities. *Annual Review of Psychology*, *66*, 547–574. doi:10.1146/annurev-psych-010814-015025

Maume, D. J. (2004). Is the glass ceiling a unique form of inequality? Evidence from a random-effects model of managerial attainment. *Work and Occupations*, *31*, 250–274. doi:10.1177/0730888404263908

Paul, M., Hamilton, D., & Darity, W., Jr. (2018, August 7). Discriminatory penalties at the intersection of race and gender in the United States. *Washington Center for Equitable Growth*. Retrieved from https://equitablegrowth.org/discriminatory-penalties-at-the-intersection-of-race-and-gender-in-the-united-states/

PayScale. (2019). *The state of the gender pay gap 2019*. Retrieved from https://archive.vn/e5vCe

Pew Research Center. (2016, June 27). *On views of race and inequality, Blacks and Whites are worlds apart*. Retrieved from http://www.pewsocialtrends.org/2016/06/27/on-views-of-race-and-inequality-blacks-and-whites-are-worlds-apart/

Prison Policy Initiative. (2019, May 14). *Policing women: Race and gender disparities in police stops, searches, and use of force*. Retrieved from https://www.prisonpolicy.org/blog/2019/05/14/policingwomen/

Rosette, A. S., Koval, C. Z., Ma, A., & Livingston, R. (2016). Race matters for women leaders: Intersectional effects on agentic deficiencies and penalties. *The Leadership Quarterly*, *27*, 429–445. doi:10.1016/j.leaqua.2016.01.008

Strand, S., Deary, I. J., & Smith, P. (2006). Sex differences in cognitive abilities test scores: A UK national picture. *British Journal of Educational Psychology*, *76*, 463–480. doi:10.1348/000709905X50906

Swets, J. A., Dawes, R. M., & Monahan, J. (2000, October). *Better decisions through science. Scientific American*. Retrieved from https://www.scientificamerican.com/article/better-decisions-through-science/

Thiem, K. C., Neel, R., Simpson, A. J., & Todd, A. R. (2019). Are Black women and girls associated with danger? Implicit racial bias at the intersection of target age and gender. *Personality and Social Psychology Bulletin*, *45*, 1427–1439. doi:10.1177/0146167219829182

Tugend, A. (2018, September 30). The effect of intersectionality in the workplace. *The New York Times*. Retrieved from https://www.nytimes.com/2018/09/30/us/the-effect-of-intersectionality-in-the-workplace.html

Urban Institute. (2017). *A matter of time: The causes and consequences of rising time served in America's prisons. Demographics*. Retrieved from https://apps.urban.org/features/long-prison-terms/demographics.html

U. S. Census Bureau. (2010). *QuickFacts: New York City, New York*. Retrieved from https://www.census.gov/quickfacts/fact/table/newyorkcitynewyork/PST040219

Vallacher, R. R., & Nowak, A. (1994). *Dynamical systems in social psychology*. San Diego, CA: Academic Press.

Vallacher, R. R., & Nowak, A. (1997). The emergence of dynamical social psychology. *Psychological Inquiry*, *8*(2), 73–99. doi:10.1207/s15327965pli0802_1

van der Linden, W. J., & Hambleton, R. K. (Eds.) (2013). *Handbook of modern item response theory*. New York, NY: Springer-Verlag.

Westwood, M., & Pinzon, J. (2008). Adolescent male health. *Paediatrics & Child Health*, *13*(1), 31–36.

Wilson, J. P., Hugenberg, K., & Rule, N. O. (2017). Racial bias in judgments of physical size and formidability: From size to threat. *Journal of Personality and Social Psychology*, *113*, 59–80. doi:10.1037/pspi0000092