Organization Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Group Size and Its Impact on Diversity-Related Perceptions and Hiring Decisions in Homogeneous Groups

Aneesh Rai, Edward H. Chang, Erika L. Kirgios, Katherine L. Milkman

To cite this article:
Aneesh Rai, Edward H. Chang, Erika L. Kirgios, Katherine L. Milkman (2024) Group Size and Its Impact on Diversity-Related Perceptions and Hiring Decisions in Homogeneous Groups. Organization Science

Published online in Articles in Advance 18 Apr 2024

. https://doi.org/10.1287/orsc.2020.14705

https://pubsonline.informs.org/journal/orsc

ORGANIZATION SCIENCE
Articles in Advance, pp. 1–26
ISSN 1047-7039 (print), ISSN 1526-5455 (online)

Group Size and Its Impact on Diversity-Related Perceptions and Hiring Decisions in Homogeneous Groups

Aneesh Rai,[a,*] Edward H. Chang,[b] Erika L. Kirgios,[c] Katherine L. Milkman[d]

[a] Department of Management and Organization, Robert H. Smith School of Business, University of Maryland, College Park, Maryland 20742; [b] Department of Negotiation, Organizations, and Markets, Harvard Business School, Harvard University, Boston, Massachusetts 02163; [c] Department of Behavioral Science, Booth School of Business, University of Chicago, Chicago, Illinois 60637; [d] Department of Operations, Information, and Decisions, The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania 19104

*Corresponding author

Contact: aneeshr@umd.edu, https://orcid.org/0000-0001-9633-2568 (AR); ehchang@hbs.edu, https://orcid.org/0000-0003-0502-2957 (EHC); erika.kirgios@chicagobooth.edu, https://orcid.org/0000-0002-7891-0499 (ELK); kmilkman@wharton.upenn.edu, https://orcid.org/0000-0002-9706-4830 (KLM)

Received: November 24, 2020
Revised: November 11, 2021; July 7, 2022; June 23, 2023; November 7, 2023; March 14, 2024
Accepted: March 14, 2024
Published Online in Articles in Advance: April 18, 2024

https://doi.org/10.1287/orsc.2020.14705

Abstract. Why do some homogeneous groups face backlash for lacking diversity, whereas others escape censure? We show that a homogeneous group's size changes how it is perceived and whether decision makers pursue greater diversity in its ranks. We theorize that people make different inferences about larger groups than smaller ones—with consequences for diversity management—due to Bayesian reasoning. This can produce sensitivity to a lack of diversity in large groups and limited sensitivity to a lack of diversity in small groups. Because each group member represents the outcome of a hiring decision, larger homogeneous groups signal a diversity problem more strongly than smaller homogeneous groups. Across three preregistered experiments ($n = 4{,}283$), we show that decision makers are more likely to diversify larger homogeneous groups than smaller ones and view larger homogeneous groups as (i) more likely to have resulted from an unfair selection process; (ii) less diverse; (iii) more likely to face diversity-related impression management concerns; and (iv) less open to the influence of newly added underrepresented members. Further, (i)–(iii) mediate the relationship between homogeneous group size and decisions to diversify. We extend our findings to S&P 1500 corporate boards, showing that larger homogeneous boards are more likely to add women or racial minorities as directors. Larger homogeneous boards are also rarer than expected, whereas smaller homogeneous boards are surprisingly abundant. This suggests that decision makers neglect homogeneity in smaller groups, while investing extra effort toward diversifying larger homogeneous groups. Our findings highlight how group size shapes diversity-related perceptions and decisions and identify mechanisms that kickstart diversification efforts.

Supplemental Material: The online supplement is available at https://doi.org/10.1287/orsc.2020.14705.

Keywords: diversity in organizations • decision making • psychological processes • archival research: extant data • laboratory research • experimental designs

In recent years, promoting organizational diversity has blossomed into a multibillion dollar industry, and many of the world's largest companies have made public commitments to furthering diversity and inclusion in their ranks (Kirkland and Bohnet 2017, Colvin 2022). However, achieving meaningful diversity has remained elusive for many organizations (Stevens 2020, Field et al. 2023). Moreover, skeptics have made compelling arguments that prevailing efforts, such as diversity statements, investments in bias trainings, and hiring chief diversity officers, represent little more than "cheap talk" (Almeida and Lordan 2020, Bunn 2023, Harmeling 2023). Indeed, homogeneous groups—those whose members belong to a single demographic category—still arise frequently in many organizations and industries (Kumar 2018, Bouvy and Mujoomdar 2019, Else 2019, Institutional Shareholder Services 2019, Larcker and Tayan 2020). There are several potential explanations for this pattern: organizations may indeed make false or exaggerated claims about their intentions to diversify (Kroeper et al. 2022), some decision makers might continue to discriminate against marginalized candidates (Bertrand and Duflo 2017), or organizations may face hurdles in recruiting candidates who add

diversity (Fernandez-Mateo and Fernandez 2016, Leslie et al. 2017). Although all of these causes may well contribute to a continuing lack of diversity, we propose another potential barrier to diversification that has largely been overlooked: decision makers may not always recognize a lack of diversity in the first place. Specifically, even if decision makers do value diversity, they may struggle under certain conditions to identify whether their groups are diverse or how others will perceive their diversity (or lack thereof).

Consider a real-world example of two law firms: in 2018, the law firm Paul, Weiss landed on the front page of *The New York Times* after announcing its new, all-White partner class, which included a single White woman (Scheiber and Eligon 2019). The announcement made Paul, Weiss the face of diversity problems in "Big Law" (Patrice 2018, Simmons 2018, Scheiber and Eligon 2019). Around the same time, another law firm, Pryor Cashman, announced an even more homogeneous partner class—one that included only White men (Pryor Cashman 2017). But Pryor Cashman didn't face any backlash. Why did these firms' announcements produce such different public reactions? There are many possible explanations, including luck or coincidence. But one potentially relevant difference is the sizes of the law firms' partner classes: Paul, Weiss named 12 new partners, whereas Pryor Cashman named just 4.

In this paper, we theorize that size—a fundamental feature of all groups—is one important feature that influences whether decision makers notice (and therefore correct) a lack of diversity in groups, even when they are homogeneous (i.e., maximally lacking in diversity). Like the media and other outside observers who censured Paul, Weiss while letting Pryor Cashman's homogeneity go unremarked, decision makers may fail to realize that homogeneous groups under their purview lack diversity (and to attribute this lack to bias) when they are smaller in size. Why is it important to understand when people recognize a lack of diversity or the potential for bias to have shaped a group's composition? Simply put, organizations cannot effectively take steps to fix a problem that has gone overlooked (Nickerson 1998, Mack 2003). To the extent that organizations wish to capitalize on the perceived business (Richard et al. 2007, Herring 2009, McKay et al. 2009; cf. Jehn et al. 1997, Thatcher et al. 2003, Leslie et al. 2023), reputational (Avery and McKay 2006), and moral (Ely and Thomas 2001, Georgeac and Rattan 2022) benefits of diversity, recognizing when a group is lacking in diversity or when bias may have influenced the group's selection processes are crucial first steps. Without such recognition, decision makers may fail to see that further efforts toward diversification are needed. However, these critical first steps have been largely understudied by diversity scholars.

Size is a fundamental and visible attribute of any organizational group or team. Prior diversity scholarship treats homogeneous groups as indistinct from one another, regardless of size (Harrison and Klein 2007, Meyer 2017). However, we draw upon a core tenet of judgment and decision-making theory—that people tend to form beliefs in an intuitive Bayesian manner (Slovic and Lichtenstein 1971, Dawes 1989, Gigerenzer and Hoffrage 1995, Moore and Healy 2008)—and propose that the size of a homogeneous group can play an important role in shaping (i) the way it is perceived, (ii) whether it will subsequently be diversified, and (iii) how aggressively decision makers pursue its diversification. According to a standard Bayesian framework, people use the information they observe to generate statistically informed guesses about the likelihood of different possibilities, which shape their beliefs and inform their actions (Slovic and Lichtenstein 1971, Green and Daniels 2021). We theorize that a logical (Bayesian) observer should interpret a larger homogeneous group as providing a stronger signal that there is a diversity problem, relative to a smaller homogeneous group (Bohnet et al. 2016). This could help explain why an all-White and primarily male 12-person partner class would generate more outrage than a partner class of 4 White men. Based on this theorizing, we posit that larger homogeneous groups will be perceived as being (i) more likely to have formed through an unfair selection process; (ii) less diverse; (iii) more likely to face diversity-related impression management concerns; and (iv) less capable of being influenced by newly added underrepresented group members than smaller homogeneous groups. We expect each of these perceptions to influence diversity-related hiring decisions, such that decision makers will be more prone to diversify larger homogeneous groups than smaller ones. We theorize that group size affects perceptions of diversity whether or not the perceiver is a member of the group being evaluated and that group size influences organizational decision makers' expectations of how their group will be perceived. These second-order perceptions may, in turn, influence personnel selection and group composition. For example, organizational decision makers who do not wish to diversify their groups may try to keep homogeneous groups small, believing that observers will then be likely to overlook their groups' homogeneity, or attribute it to chance rather than bias.

Understanding when people are (or aren't) motivated to diversify groups has important implications for organizations. Research on intergroup contact suggests that interactions with outgroup members can help reduce prejudice, and other work finds that exposure to counterstereotypical exemplars can reduce implicit bias and help people overcome stereotypes (Lai et al. 2013, Finnegan et al. 2015, Robertson and Weiss

2017, Paluck et al. 2019, Corno et al. 2022). Because diversifying homogeneous groups creates opportunities for intergroup contact and can facilitate exposure to counterstereotypical exemplars, better understanding the forces that lead homogeneous groups to diversify may be particularly important for reducing bias and accelerating diversification efforts in organizations more broadly. In addition, previous work suggests that homogeneity itself can exert distinct and negative effects on groups, with implications for prejudice, discrimination, and intergroup conflict (Apfelbaum et al. 2014). Thus, understanding what initiates the diversification of homogeneous groups holds practical significance.

Finally, by incorporating a Bayesian reasoning model into the diversity literature, we illuminate how an understudied but central feature of all groups—their size—affects how decision makers perceive and work to alter their organizations' diversity levels. Group size is often neglected in current definitions of diversity, which treat homogeneous groups of varying sizes as definitionally identical (Harrison and Klein 2007, Meyer 2017). It is also often treated as an ancillary control variable in diversity scholarship (e.g., Riordan and Shore 1997, Chatman and O'Reilly 2004, van Knippenberg et al. 2007, Hentschel et al. 2013, Chang et al. 2019). We push back against these notions by suggesting that diversity scholars should pay closer attention to the size of a group and how it may, in concert with other features of the group, influence diversity-related perceptions and behavior. Through the lens of Bayesian reasoning, we identify how group size interacts with group composition to predict the inferences people make about a group's selection process, diversity levels, and impression management concerns—all of which, in turn, affect people's motivation to diversify the group.

Homogeneity, Group Size, and Diversity-Related Perceptions

We propose that a homogeneous group's size can shape the way it is perceived, as well as whether efforts to diversify it will be initiated, offering insight into motives that can more generally propel diversification efforts. Why might group size matter in this way? A core tenet of judgment and decision-making theory is that people tend to form beliefs in a logical, Bayesian manner, acting like intuitive statisticians even when making automatic judgments (Slovic and Lichtenstein 1971, Dawes 1989, Gigerenzer and Hoffrage 1995, Kersten et al. 2004, Charness and Levin 2005, Griffiths and Tenenbaum 2006, Moore and Small 2008, Grieco and Hogarth 2009, Green and Daniels 2021). For example, Green and Daniels (2021) show that baseball umpires making split-second decisions seem to respond with a Bayesian instinct when calling pitches by integrating the imperfect information they observe about a pitch's location with their expectations about the pitch's most likely location.

We propose that people also use Bayesian reasoning when forming their perceptions of a group. A group of people in an organization can be conceptualized as the result of a series of personnel selection decisions, with each individual group member representing the outcome of one such decision. Larger groups then provide a stronger signal for making inferences about the group, such as whether the group is diverse or not, relative to smaller groups. In other words, because larger groups can be thought of as providing a larger "sample size," evaluators have more information to assess whether the group is representative of the underlying population when judging its diversity (or lack thereof).

For example, imagine two people—John and Jane—are each evaluating the diversity of their teams, both of which were created by selecting members from a population equally composed of Computer Science (CS) majors and Math majors. John's group has four members, whereas Jane's group has eight members. Each one knows that their group could have been created through one of two hiring processes: Process A, which is biased in favor of CS majors (so around 90% of hires are CS majors and around 10% are Math majors), or Process B, which doesn't favor any group (so around 50% of hires are CS majors and around 50% are Math majors). Before seeing their groups, John and Jane may believe it is equally likely that they have been assigned a group created through Process A or Process B. However, if John learned that his four-person group consisted entirely of CS majors, it would be appropriate for him to suspect that his group was more likely created through Process A than Process B. On the other hand, if Jane saw that her eight-person group consisted entirely of CS majors, it would be appropriate for her to have an *even stronger suspicion* than John that her group was selected through Process A rather than Process B because the larger size of the team she is assessing provides a larger sample from which to draw conclusions about possible bias in the selection system. In other words, Bayesian reasoning should lead people to form different judgments about larger homogeneous groups than smaller ones. Put another way, seeing eight consecutive coin flips land on heads offers a stronger signal that one is observing a biased coin than seeing the same coin land on heads just four consecutive times.

Consistent with this idea, there is some prior empirical evidence that individuals form beliefs and judgments about groups of different sizes as a Bayesian model of reasoning would predict. For example, Kerr (1989) found that group size affected people's perceptions of collective efficacy in public goods games. The author theorized that each individual added to a group increases the perceived risk of diffusion of responsibility, free-riding, and

lack of individual accountability. Accordingly, larger groups are perceived as having less collective efficacy than smaller groups. We expect a similar process to play out with respect to diversity-related perceptions of a group: larger groups, compared with smaller groups, provide evaluators with a stronger signal about whether and how sharply the group's composition deviates from the composition of the underlying population from which group members were drawn. This should influence the diversity-related perceptions that evaluators form (whether evaluators are members of the group themselves). Further, we propose that these perceptions can be expected to shape key decisions about who else should be added to the group.

Although our theorizing focuses on homogeneous groups, with implications for understanding what will kick-start diversification efforts in such groups, it can naturally be extended to the study of nonhomogeneous groups as well. Larger groups—whether they are homogeneous or not—provide more information to evaluators. However, in the case of diverse groups, for example, a larger group size provides a stronger signal about the group's diversity (rather than its lack thereof). In our general discussion, we elaborate on the predictions about diverse groups that result from our theorizing, and we discuss data from a supplemental study that supports those predictions.

Perceptions of a Homogeneous Group's Selection Process

People tend to perceive an organization's selection process as "fair" when it shows no bias against particular applicants and as "unfair" when bias exists (Leventhal et al. 1980, Folger and Bies 1989, Gilliland 1993, Blader 2007). By definition, homogeneous groups only contain members of a single demographic category. Thus, homogeneous groups may be perceived as the product of a discriminatory selection process involving bias against members of unrepresented demographic categories (Gilliland 1993, Harris et al. 2004, Patterson and Zibarras 2011). If people act like intuitive Bayesians, then because larger homogeneous groups provide a stronger signal than smaller ones about the selection process of group members, people should infer that larger homogeneous groups were more likely formed by biased—and therefore unfair—selection processes than smaller homogeneous groups (Leventhal et al. 1980, Blader 2007, Holt and Smith 2009).

For example, an all-male (or all-White) team might be seen as providing potential evidence of a hiring process that is biased against women (or racial minorities; Gilliland 1993, Patterson and Zibarras 2011). This belief should be stronger among Bayesians the larger the all-male (or all-White) team is because, for larger teams, the probability of not selecting a woman (or racial minority) for *any* position by chance—rather than due to bias—is smaller (Maxwell et al. 2008). Thus, we hypothesize the following:

Hypothesis 1(a). *Larger homogeneous groups will be perceived as more likely to stem from unfair selection processes than smaller homogeneous groups.*

Perceptions of a Homogeneous Group's Diversity Levels

Objective diversity focuses on the presence of differences within a group, whereas perceived diversity is defined by whether people recognize and are aware of those differences (Harrison and Klein 2007, Bauman et al. 2014, Shemla et al. 2016). Perceived diversity is also driven by subjective factors, such as motivated reasoning and self-serving biases (Unzueta and Binning 2012, Unzueta et al. 2012, Bauman et al. 2014). Thus, different evaluators can perceive the same group's diversity very differently. Divergent perceptions of a group's diversity are particularly likely to arise under conditions of ambiguity, where there is more leeway for subjective factors to distort judgments (Kunda 1990).

However, in the context of homogeneous groups, as the size of the group increases, there is less ambiguity surrounding a group's lack of diversity. This is because larger groups provide a larger sample size, and, thus, for Bayesian observers, larger (as compared with smaller) homogeneous groups provide stronger signals that the group's composition lacks diversity (Holt and Smith 2009). The stronger signals mean that Bayesian observers can be more confident in their judgment about larger homogeneous groups lacking diversity, relative to smaller ones. Even for observers who are motivated to construe a group as diverse, reductions in ambiguity of this type should make it harder to perceive a homogeneous group as diverse (Dunning et al. 1989, Kunda 1990). Thus, in general, Bayesian evaluators should be more certain that a larger homogeneous group is, indeed, lacking in diversity than a smaller homogeneous group. This leads us to hypothesize the following:

Hypothesis 1(b). *Larger homogeneous groups will be perceived as less diverse than smaller homogeneous groups.*

Perceptions of a Homogeneous Group's Diversity-Related Impression Management Concerns

Impression management describes efforts by individuals, groups, or organizations to ensure that they are perceived in a positive light (Elsbach and Sutton 1992, Highhouse et al. 2009). Impression management concerns arise when there is reason to worry that outside parties may have a negative perception of a group or organization that could harm it in some way (Dutton et al. 1994).

Some groups and organizations pursue greater diversity due to impression management concerns. For example, Chang et al. (2019) offer evidence that organizations try to increase the diversity of highly visible groups in response to impression management concerns and note that groups lacking in diversity can face negative reputational consequences (like the backlash faced by Paul, Weiss in our opening example). Whereas their research focuses on how social norms influence organizations to achieve certain diversity "thresholds," we focus on how a homogeneous group's size influences choices about whether to begin diversifying that group at all.

We propose that larger homogeneous groups are viewed as more likely to face diversity-related impression management concerns than smaller ones. This prediction is based on simulation theory (Gallese and Goldman 1998, Shanton and Goldman 2010), which suggests that individuals predict others' thoughts and mental states by using *their own* thoughts and mental states as a template to simulate what others might be thinking. Thus, if evaluators form diversity-related judgments about a homogeneous group using Bayesian reasoning, as we propose, then they should expect outside observers to judge the group similarly. Specifically, they should intuit that observers will view a larger homogeneous group as less diverse and more likely to have been formed by a biased selection process than a smaller homogeneous group. Importantly, they should also be concerned about outside observers arriving at similar conclusions. This would be more likely to generate diversity-related impression management concerns for larger homogeneous groups than smaller ones.

Hypothesis 1(c). *Larger homogeneous groups will be perceived as more likely to face diversity-related impression management concerns than smaller homogeneous groups.*

Perceptions of the Influence Wielded by New Group Members Who Add Diversity in Homogeneous Groups

An individual group member's influence within a group and over group outcomes generally diminishes in larger groups relative to smaller groups (Latané et al. 1979, Kerr 1989, Carron and Spink 1995, Spreitzer 1995). Although this is true for any newly added group member, demographically underrepresented group members added to previously homogeneous groups may be particularly likely to wield less influence as group size increases. This is because an underrepresented member added to a group dilutes the group's homogeneity less in larger than in smaller groups, and research on tokenism suggests that the negative consequences of being underrepresented in a group are exacerbated when the group is more "skewed"—that is, when the demographic minorities in the group are less well-represented (Kanter 1977). For example, a woman added to an all-male group of four creates a group whose composition is 20% female, but if she, instead, joins an all-male group of nine, the new group will be only 10% female.

The potential negative consequences of being underrepresented in a group include feelings of exclusion or ostracism, identity-based discrimination, and pressure to behave according to identity-based stereotypes (Kanter 1977, MacCorquodale and Jensen 1993, Mehra et al. 1998, Schaffer and Riordan 2013, Watkins et al. 2019). Past research has shown a universal awareness that tokens may face negative consequences (MacCorquodale and Jensen 1993). Therefore, as the size of a homogeneous group increases, meaning that a newly added underrepresented member would be more outnumbered, Bayesian observers should form a stronger belief that the new underrepresented member faces a risk of experiencing the negative effects of tokenization. We therefore predict that for larger homogeneous groups, a newly added member who is demographically underrepresented will be expected to wield less influence.

Hypothesis 1(d). *Larger homogeneous groups will be perceived as allowing newly added demographically underrepresented members to wield less influence than smaller homogeneous groups.*

Effects of Homogeneous Group Size on Diversity-Related Hiring Decisions

Past research has demonstrated that most people dislike inequity and unfairness and try to rectify perceived injustices they identify (Fehr and Schmidt 1999, Fehr and Fischbacher 2004, Hegtvedt et al. 2009, Lotz et al. 2011). This suggests that the more likely people presume it is that a group was formed through an unfair selection process that denied opportunities to members of underrepresented groups, the more likely they will be to try to diversify the group in an attempt to correct this.

Further, people are generally more likely to diversify groups that are seen as insufficiently diverse, whether due to an organization's failure to represent its customer base, a mismatch in representation across an organization's hierarchy, or some other failure to meet diversity goals (Ely and Thomas 2001, Ng and Wiesner 2007, Avery et al. 2012, Koch et al. 2015, Beaurain and Masclet 2016). This suggests that the less diverse decision makers perceive a group to be, the more likely they will be to diversify it.

In addition, organizations have been shown to preemptively use impression management strategies when anticipating negative reputational consequences (Elsbach et al. 1998, Graffin et al. 2011). In other words, if organizations *anticipate* that their actions might generate negative perceptions, they preemptively engage

in impression management techniques to protect their public image. Therefore, we expect that as decision makers perceive a group to be more at risk for triggering negative diversity-related impressions, they will be more likely to diversify the group in an effort to avoid the potential negative reputational consequences of lacking diversity.

Finally, people may prefer to add demographic minorities to homogeneous groups when they believe those minorities will be less influential. Despite the potential benefits of diversity, many continue to view diversity as costly (Mannix and Neale 2005, van Knippenberg and Schippers 2007) and may hold biased attitudes about the competence of people from underrepresented groups, which can lead to discriminatory hiring decisions (Heilman 1983, Lyness and Heilman 2006, Petsko and Rosette 2023). Further, diversity can increase actual or perceived conflict within groups (Jehn et al. 1997, Thatcher et al. 2003, Lount et al. 2015). As a result, people may see diversifying a homogeneous group as having accompanying costs, despite the benefits for fairness and impression management. However, if people expect demographic minorities to be less influential in larger groups, then adding an underrepresented group member to a larger team may be seen as an effective way to capture the potential benefits of an underrepresented group member's presence while minimizing the perceived accompanying costs. Given that economic theory predicts the demand for a resource has an inverse relationship with its cost (Mankiw 2020), we would expect decision makers to be more willing to diversify larger homogeneous groups than smaller ones.

Together, this leads us to make the following predictions:

Hypothesis 2. *The likelihood of selecting an underrepresented candidate for a position in a homogeneous group will be higher for larger groups than smaller groups.*

Hypothesis 3(a). *The effect of group size on the likelihood of selecting an underrepresented candidate to join a homogeneous group will be mediated by the perceived fairness of the homogeneous group's selection process.*

Hypothesis 3(b). *The effect of group size on the likelihood of selecting an underrepresented candidate to join a homogeneous group will be mediated by the homogeneous group's perceived diversity.*

Hypothesis 3(c). *The effect of group size on the likelihood of selecting an underrepresented candidate to join a homogeneous group will be mediated by the perceived diversity-related impression management concerns faced by the homogeneous group.*

Hypothesis 3(d). *The effect of group size on the likelihood of selecting an underrepresented candidate to join a homogeneous group will be mediated by the perceived influence wielded by the underrepresented candidate if they were added to the homogeneous group.*

If decision makers are especially motivated to diversify larger homogeneous groups by adding underrepresented members (e.g., adding a woman to an all-male group or a non-White person to an all-White group), then as a downstream consequence of decision makers exerting extra effort to diversify larger homogeneous groups, larger homogeneous groups should be underrepresented relative to chance expectations. In other words, if decision makers disproportionately react to a lack of diversity in larger homogeneous groups, but not smaller ones, and if they respond by working especially hard to diversify larger homogeneous groups, then we should see a relative scarcity of large homogeneous groups compared with expectations.

Hypothesis 4. *As group size increases, the likelihood of observing homogeneity will decrease significantly more than predicted by chance.*

Overview of Studies

In this paper, we present evidence from three experiments and a field study testing our hypotheses. Across all four studies, we examine how homogeneous group size influences the likelihood of selecting an underrepresented candidate (Hypothesis 2). In Study 3, we also test the mechanisms that we theorize underlie the relationship between homogeneous group size and diversification decisions (Hypotheses 1 and 3). In Study 4, we extend our online studies to the field and present data to establish the external validity of our theory about homogeneous groups. Study 4 also tests our hypothesis that larger homogeneous groups should be particularly underrepresented "in the wild," following our theory that decision makers will exert more effort to avoid homogeneity in large groups than small ones (Hypothesis 4). Data for Studies 1–3 as well as analysis code for all studies can be found on OSF: https://osf.io/JZR6D/.

Study 1: How the Size of an All-White Group Affects the Decision to Hire a Black Candidate

In Study 1, we test whether people are more likely to hire a Black employee to join an all-White group as its size increases (Hypothesis 2).

Method

Participants. We recruited 600 participants through Amazon Mechanical Turk (50% identified as men, 75% identified as White). Participants were paid $0.45 for completing a survey they were told would take approximately three minutes. This study was preregistered on AsPredicted.org (https://aspredicted.org/wc5wz.pdf).

Procedures. Participants were asked to imagine that they were hiring a consultant to join a work group at their consulting company. They were given a job description for the work group's open position and were then shown the names and faces (taken from the Chicago Face Database; Ma et al. 2015) of the current members of the work group. All participants were shown an all-male, all-White group. We randomly varied the size of the work group so that it included one to eight people. That is, participants in *condition 1* saw a work group including one White man, participants in *condition 2* saw a work group including two White men, and so on, up to participants in *condition 8*, who saw a work group including eight White men.

Next, participants were shown a set of three candidates—two White men and one Black man—and asked to hire one of them to join the work group.[1] Participants were provided with the candidates' pictures (again taken from the Chicago Face Database; Ma et al. 2015) and qualifications, which included the candidates' education, years of experience, and most recent job. We stimulus-sampled candidate pictures and qualifications across participants to ensure that our effects were not driven by the particular stimuli chosen.[2]

After making a hiring decision, participants answered manipulation check questions asking them to recall the size of the original work group displayed to them, as well as the number of Black consultants in the original work group (the correct answer was zero).[3] Finally, participants reported their gender and race. Study materials are available in the online supplement.[4]

Results and Discussion

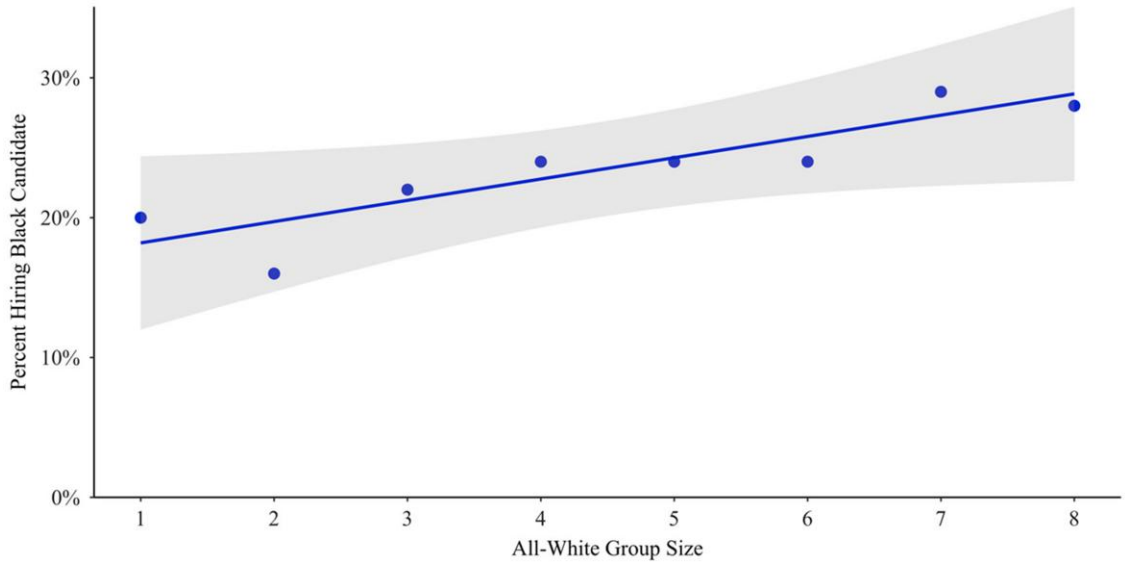
Confirming that our manipulation was effective, participants reported seeing a larger work group when we showed them a larger group ($b = 0.95$, $p < 0.001$), and 77% of participants accurately recalled that the original work group did not include any Black employees.[5] A correlation matrix of all variables collected in this study is available in the online supplement (Table S1).

Our dependent variable of interest was a binary indicator for whether participants chose to hire the Black male candidate. Following our preregistered analysis plan, we ran an ordinary least squares (OLS) regression with robust standard errors to predict whether a Black male candidate was hired. We relied on a linear model because it yields easily interpretable coefficients (Gomila 2021), but also report results from a logistic regression model, which are extremely similar.

Our only independent variable was the size of the all-White work group shown to participants, which varied from including one person to including eight people. Participants were significantly more likely to hire a Black male candidate the larger the size of the all-White work group ($b = 0.015$, $p = 0.044$; see Figure 1 and online supplement, Table S2), providing support for Hypothesis 2. We replicate this result when we analyze our data using a logistic regression instead of an OLS regression ($b = 0.085$, $p = 0.045$; see online supplement, Table S2).[6]

Study 1 provides initial empirical support for our hypothesis that decision makers are more likely to diversify larger racially homogeneous groups than smaller

Figure 1. (Color online) Likelihood of Hiring a Black Candidate as a Function of the Randomly Assigned Size of the All-White Group He Would Be Joining in Study 1



Notes. Dots represent the raw proportions of participants hiring the Black candidate across conditions. The line represents the fitted linear regression line from the analysis described in-text. The shaded region represents a 95% confidence interval.

ones. Study S1 in our online supplement presents a conceptual replication in the context of gender diversity, showing that decision makers are also more likely to add women to larger all-male groups than smaller ones.

Study 2: How the Size of an All-Male Group Affects Real Decisions to Recommend Women

In Study 2, we extend the results of Study 1 to a setting where participants make real (rather than hypothetical) decisions when offered the opportunity to add gender diversity to a homogeneous group.

Method

Participants. We recruited 2,373 college-educated participants through Prolific Academic (42% identified as men).[7] Participants were paid $0.60 to complete a roughly four-minute survey. This study was preregistered on AsPredicted.org (https://aspredicted.org/jb4fy.pdf).

Procedures. Participants were truthfully told that an East Coast business school was seeking recommendations for professors to invite as speakers for an online seminar series that was intended to appeal to a broad audience. Participants were given a description of the seminar series, and they were shown a group of professors who had already been selected for inclusion in this seminar series. Specifically, for each professor already included in the seminar series, participants were shown their name, face (drawn from the professor's website), institution, years of experience (taken from the professor's CV), area of expertise, and seminar topic. All participants were shown a group of all-male, all-White speakers. We randomly varied whether participants were shown two (*size two condition*) or eight (*size eight condition*) White male speakers who had already been invited to the seminar. Thus, we manipulated whether participants were assigned to see a smaller or larger homogeneous group.

Next, participants were shown a set of three potential speakers: two White male professors and one White female professor. They were asked to recommend one person to add to the seminar series. Participants were provided with the potential speakers' names, faces, institutions, years of experience, areas of expertise, and a potential topics they could speak on (all information was drawn from actual candidate speakers' websites and CVs). As with our previous studies, we stimulus-sampled candidates across participants.[8] After concluding data collection, we shared all recommendations with the organizers of the seminar series.

After making their speaker recommendation, participants answered manipulation check questions that asked them to recall the size of the original group of included speakers, as well as the number of women in the original group (the correct answer was zero). Finally, participants reported their gender. Complete study materials are available in the online supplement.

Results and Discussion

Confirming that our manipulation was again effective, participants reported seeing a larger original group of speakers in the *size eight condition* than the *size two condition* ($b = 5.44$, $p < 0.001$), and 73% of participants accurately recalled that the original group of speakers did not include any women.[9] A correlation matrix of all variables collected in this study is available in the online supplement (Table S3).

Our dependent variable of interest was a binary indicator for whether participants chose to recommend that the female professor be added as a speaker. Following our preregistered analysis plan, we ran a two-tailed, two-sample proportions test comparing how many participants recommended the female professor across conditions. Participants were significantly more likely to recommend the female professor in the *size eight condition* (43.9% did so) than in the *size two condition* (35.8% did so, $p < 0.001$). This provides further support for Hypothesis 2 in a setting involving real recommendation decisions.

Study 3: The Mediating Role of Diversity-Related Perceptions

In Study 3, we test Hypotheses 1, (a)–(d) and 3, (a)–(d), exploring the mechanisms responsible for the relationship between homogeneous group size and selection decisions documented in Studies 1 and 2.

Method

Participants. We recruited 1,310 participants through Amazon Mechanical Turk (45% identified as men). Participants were paid $0.50 for completing a roughly four-minute survey. This study was preregistered on AsPredicted.org (https://aspredicted.org/e7yk7.pdf).

Procedures. Participants were asked to imagine they were a hiring manager at a technology company tasked with hiring a software engineer to join an Innovation Team. Participants were given a job description for the Innovation Team's open position, and they were then shown the names and faces (taken from the Chicago Face Database; Ma et al. 2015) of the current members of the Innovation Team. All participants were shown an all-male, all-White current Innovation Team. As in Study 1, we randomly varied the size of the current Innovation Team from having as few as one to having as many as eight (White male) members.

Next, participants were shown a set of three candidates—two White men and one White woman—and asked to select one of them to join the Innovation Team. Participants were provided with the candidates' pictures (taken from the Chicago Face Database; Ma et al. 2015) and qualifications, which included the candidates' years of work experience and most recent job titles. We stimulus-sampled candidate pictures and qualifications across participants.[10]

After participants made their hypothetical hiring decision, they answered a series of questions measuring our hypothesized mediators. All scale items used in this study are available in the online supplement. All items across all scales were measured using 7-point Likert scales, where 1 was defined as "Strongly Disagree" and 7 was defined as "Strongly Agree."

Perceived Fairness of the Original Group's Selection Procedure. To measure the perceived fairness of the process used to select the members of the original Innovation Team, we adapted one question from Sweeney and McFarlin's (1997) procedural justice scale ("The procedures used to evaluate and select the members of the original Innovation Team were likely fair and objective.").

Perceived Diversity of the Original Group. To measure the perceived diversity of the original Innovation Team, we used a three-item scale adapted from Unzueta and Binning's (2012) perceived diversity scale (Cronbach's $\alpha = 0.88$; e.g., "The original Innovation Team had a high degree of gender diversity.").

Diversity-Related Impression Management Concerns Faced by the Original Group. To measure the diversity-related impression management concerns relating to the original Innovation Team, we used a two-item scale adapted from the Chng et al. (2015) image concerns scale (Cronbach's $\alpha = 0.92$; e.g., "The tech company should be worried about how key stakeholders will perceive the diversity of the original Innovation Team.").

Potential for a New Addition to Influence the Group. To measure the perceived influence a woman would wield if she were added to the original Innovation Team, we used a three-item scale adapted from Spreitzer's (1995) impact scale (Cronbach's $\alpha = 0.83$; e.g., "If a woman were added to the original Innovation Team, she would have a large impact on the group's work.").

Multi-item scales were averaged across items. After responding to our mediator questions, participants answered manipulation check questions asking them to recall the size of the original Innovation Team, as well as the number of women on the original Innovation Team. Finally, participants reported their gender. Study materials are available in the online supplement.

Results and Discussion

Confirming that our manipulation was effective, participants reported seeing a larger Innovation Team when we showed them a larger one ($b = 0.86$, $p < 0.001$). In addition, 87% of participants accurately reported seeing zero women in the Innovation Team.[11] A correlation matrix of all variables collected in this study is available in the online supplement (Table S4).

First, following our preregistered analysis plan, we ran an OLS regression with robust standard errors to predict whether a female candidate was hired. The dependent variable was a binary indicator for whether a female candidate was hired. Our only independent variable was the size of the all-male Innovation Team shown to participants, which varied from one to eight. Consistent with Hypothesis 2, participants were significantly more likely to select a female candidate for larger all-male Innovation Teams ($b = 0.027$, $p < 0.001$; see online supplement, Table S5). We replicate this result when we analyze the data using a logistic regression instead of an OLS regression ($b = 0.148$, $p < 0.001$; see online supplement, Table S5).[12]

Next, we tested whether group size had an effect on participants' diversity-related perceptions of the homogeneous group (Hypotheses 1(a)–1(d)). To test Hypotheses 1(a)–1(d), we ran four separate OLS regressions with all-male Innovation Team size as the independent variable. The dependent variables in these four regressions were: (i) the perceived fairness of the Innovation Team's selection process (to test Hypothesis 1(a)); (ii) the perceived diversity of the Innovation team (to test Hypothesis 1(b)); (iii) the perceived diversity-related impression management concerns faced by the Innovation Team (to test Hypothesis 1(c)); and (iv) the perceived influence a woman would wield if she was added to the Innovation Team (to test Hypothesis 1(d)). We find that participants judged the group's selection process as more unfair for larger all-male Innovation Teams than smaller all-male Innovation Teams ($b = -0.131$, $p < 0.001$), supporting Hypothesis 1(a). Participants also perceived larger all-male Innovation Teams to be less diverse than smaller all-male Innovation Teams ($b = -0.127$, $p < 0.001$), supporting Hypothesis 1(b). Supporting Hypothesis 1(c), participants believed the larger all-male Innovation teams were more likely to face diversity-related impression management concerns than the smaller all-male Innovation teams ($b = 0.145$, $p < 0.001$). Finally, consistent with Hypothesis 1(d), participants believed a woman added to an all-male Innovation Team would wield less influence in larger teams than in smaller ones ($b = -0.038$, $p = 0.008$). See online supplement, Table S6 for detailed regression results.

Finally, we tested whether our hypothesized mechanisms mediated the effect of an all-male group's size on participants' likelihood of hiring a female candidate to join the group (Hypotheses 3(a)–3(d)). Following our

preregistration, we ran a 5,000-sample bootstrapped multiple mediation model (Preacher and Hayes 2008) including measures of all four hypothesized mediators. We find that all four hypothesized mediators—perceived fairness of the original all-male Innovation Team's selection process ($b = 0.006$, $p < 0.001$; 95% CI: [0.004, 0.009]), perceived diversity of the original Innovation Team ($b = 0.007$, $p < 0.001$; 95% CI: [0.003, 0.008]), diversity-related impression management concerns faced by the original Innovation Team ($b = 0.009$, $p < 0.001$; 95% CI: [0.006, 0.012]), and perceived influence a woman would wield if added to the original Innovation Team ($b = -0.002$, $p = 0.016$; 95% CI: [−0.004, −0.0004])—predicted the relationship between the randomly assigned size of the original Innovation Team and participants' likelihood of hiring a female candidate to join the team, providing support for Hypotheses 3(a)–3(d). The full results from this mediation model can be found in the online supplement (Figure S1). We also tested each mediator individually using separate mediation models, and those analyses provide convergent support for our hypotheses (see online supplement, Figure S2).

Notably, the association we measured between the anticipated influence a female candidate would exert if added to an Innovation Team and participants' propensity to hire a female candidate ran counter to our hypothesizing (Hypothesis 3(d)). We predicted that a woman added to an all-male group would be expected to wield less influence in larger all-male groups, and this is what we found. However, contrary to our theorizing, participants were more likely to hire a woman into an all-male group if they thought she would wield *more* influence ($b = 0.077$, $p < 0.001$). This suggests that participants considered adding a woman's perspective to an all-male group to be a benefit rather than a "risk" or "cost" to be avoided. This may be due to the potential benefits of diversity on outcomes such as organizational performance, financial performance, creativity, and information-sharing within groups (Phillips 2003, Phillips and Loyd 2006, Richard et al. 2007, Herring 2009, McKay et al. 2009, Sinaceur et al. 2010). Our finding highlights one possible way in which larger homogeneous groups are *less* attractive targets for diversification than smaller ones—larger homogeneous groups may be seen as less likely to reap the potential benefits of diversification.

In addition to conducting our preregistered analyses, as a robustness check, we reran our analyses excluding the experimental condition involving an Innovation Team with just one White male member (because several of our mediator measures asked about the original Innovation Team, and, arguably, a single person cannot be evaluated as a team). When we do this (see online supplement, Figure S3), our results are unchanged, with one exception: we no longer find support for Hypothesis 1(d). Specifically, the negative effect of the all-male group's size on the perceived influence a woman would wield if added becomes insignificant ($b = -0.010$, $p = 0.664$).

Study 4: Group Homogeneity and Size on U.S. Corporate Boards

In Study 4, we move from online experiments to the field, testing the predictions of our theory in an important organizational context: U.S. corporate boards in the S&P 1500. Although corporate boards represent just one possible context where our theorizing about homogeneous groups may apply, these groups exert considerable influence, overseeing companies in the S&P 1500 that represent roughly 90% of the total U.S. stock market capitalization (S&P Dow Jones Indices 2019). Corporate boards in the United States are also often subject to external pressures relating to their lack of diversity (McGreevy 2018, Green 2020), which means that board members should care about how their group might be perceived. We join a growing literature on understanding diversity in corporate boards in light of the important consequences it can have in this context (Dezső et al. 2016, Tinsley et al. 2017, Chang et al. 2019, Lawson et al. 2022).

Study 4A: Board Size and the Likelihood of Adding Underrepresented Group Members

To test Hypothesis 2 in a field setting, in Study 4A, we first examine whether larger, all-male S&P 1500 corporate boards are more likely than smaller all-male boards to diversify by adding new, female directors. We then examine whether all-White S&P 1500 corporate boards are more likely than smaller all-White boards to diversify by adding new, non-White directors.

Method

Data. The data for these analyses come from Institutional Shareholder Services (ISS) Director Data. ISS Director Data includes information on individual members of the boards of directors of companies in the S&P Composite 1500, including each director's name, gender, and race/ethnicity.[13] For this analysis, we used data from 2007—the earliest year for which data were available in a consistent format—through 2018, the most recent year of data available to us as of July 21, 2019, when we first accessed the ISS database. To test Hypothesis 2, which claims that larger homogeneous boards will be more likely to add an underrepresented member than smaller homogeneous boards, we restricted our data set to years in which a company had an all-male board. For each all-male board, we have information about its composition during year t and year $t + 1$. This allows us to use the all-male board's size in year t to

predict whether it added any women in year $t+1$. Thus, our unit of observation was a board-year because an individual company could contribute multiple observations to the data set. This left us with 3,593 board-year observations (across 836 unique corporate boards) to analyze.

Control Variables. We included a battery of control variables in our analyses to help rule out alternative explanations for our findings.[14] We included the following financial performance indicators and firm metrics to alleviate concerns about firm size or performance as alternative explanations for our findings: a firm's market capitalization in a given year,[15] a firm's return-on-assets in a given year, a firm's logged total assets in a given year, a firm's one-year total-shareholder-returns in a given year, a firm's market-to-book ratio in a given year, a firm's Tobin's Q in a given year, a firm's number of employees in a given year, and the logged compensation of a firm's CEO in a given year.

We included the following controls to account for the possibility that our findings reflect differences in corporate governance among firms with larger versus smaller boards: a firm's level of institutional ownership in a given year measured using percentage of shares owned by institutions (Brown et al. 2017), the proportion of a firm's board members who are independent in a given year, a firm's board's classification (i.e., whether its members are elected to terms that expire in different years), and whether a firm's board has a majority voting system.

To account for the possibility that other dimensions of board diversity influence decision making, which could covary with board size, we included a firm's board's age diversity, as measured by the standard deviation of all members' ages in a given year (Triana et al. 2014), and a firm's board's tenure diversity, as measured by the standard deviation of all members' tenure on the board in a given year (Triana et al. 2014). We also included fixed effects for calendar year in our analyses to account for time trends. Finally, we included fixed effects for the industry that a firm belongs to in order to alleviate concerns that specific industries might also have larger or smaller boards, on average, which could be driving the effects of interest.[16]

Following the recommendation of Becker (2005), we present our analyses both with and without the aforementioned controls. In case of missing data for any of our control variables, we added an additional "missing" category for each categorical variable, and we used mean imputation and included binary indicators for all missing continuous variables. We winsorized all continuous control variables that were not logged at the 0.5th and 99.5th percentiles to prevent outliers from exerting undue influence on our analyses. However, our results are essentially identical even if we do not winsorize any control variables.

Analysis Strategy. We estimated OLS and logistic regressions to predict whether an all-male board added a woman to its ranks in a given year (thus eliminating its homogeneity). We clustered standard errors at the board level to account for nonindependence of data due to repeated observations for firms across years. Our binary dependent measure took on a value of 1 if the all-male board added at least one woman between year t and year $t+1$, and it took on a value of 0 otherwise. Our main predictor was the size of the all-male board.

Results

Summary Statistics. A correlation matrix of all variables used in this study is available in the online supplement (Table S7). Of the 3,593 all-male board-year observations, 564 (15.7%) involved boards that added at least one woman in the following year. Boards ranged in size from 3 to 18, with a median size of 7. We winsorized board size at the 0.5th and 99.5th percentiles to prevent outliers from exerting undue influence on our analyses. The winsorized board size distribution ranged from 4 to 14. Robustness checks where we do not winsorize the size distribution are reported in the online supplement (Table S8) and yield similar results to those reported below.

Are Larger All-Male Boards More Likely to Add Women? As Table 1, Model (1) shows (using an OLS regression), the size of an all-male board is a significant and positive predictor of whether a woman was added to the board in the following year ($b = 0.019$, $p < 0.001$). Specifically, there was an average increase of 1.9 percentage points in the likelihood that a given all-male board would diversify in a given year associated with each additional board member. As Table 1, Model (2) shows, this result is consistent when we use a logistic regression ($b = 0.141$, $p < 0.001$, $OR = 1.15$). Thus, larger all-male boards were more likely to add a woman to their ranks than smaller all-male boards, consistent with Hypothesis 2. Figure 2 shows that the likelihood of adding a woman increases across the range of possible board sizes. As shown in Table 1, Model (3), when we add our control variables, we find that our predictor for board size is still positive and significant ($b = 0.009$, $p = 0.039$), corresponding to about a 0.9 percentage point increase in the likelihood of an all-male board diversifying for each additional member. As Table 1, Model (4) shows, this result is consistent, though marginally significant, when we use a logistic regression ($b = 0.059$, $p = 0.087$, $OR = 1.06$).

As a robustness check, we reran our regression model, but omitted any control variables that had an

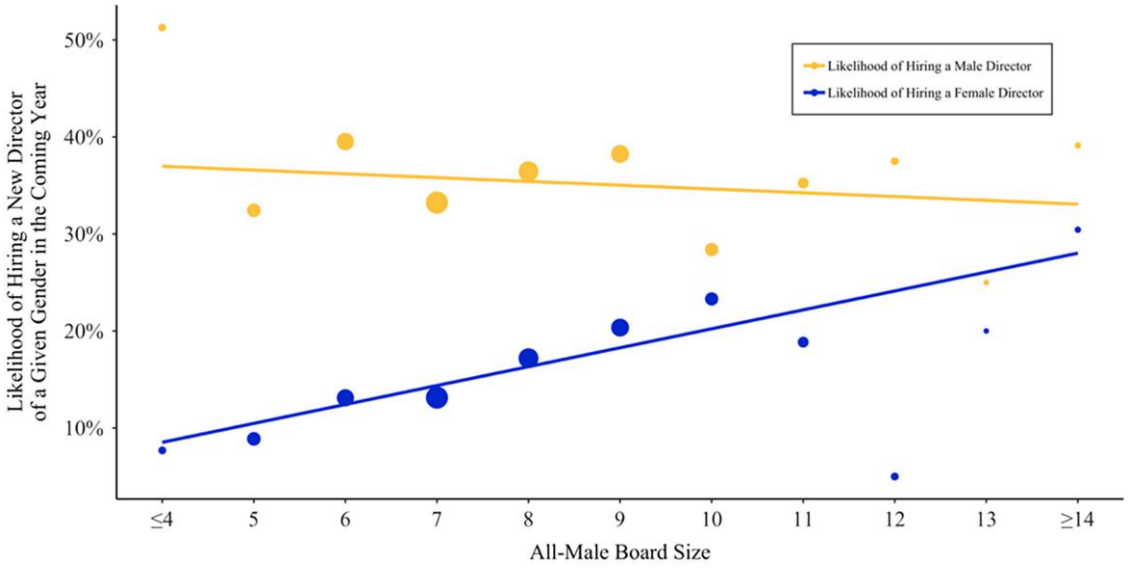
Table 1. Larger All-Male Boards Are More Likely to Add Women to Their Ranks

	Dependent variable: *Woman Added to All-Male Board*					
Variable	Model (1): OLS regression	Model (2): Logistic regression	Model (3): OLS regression	Model (4): Logistic regression	Model (5): OLS regression	Model (6): Logistic regression
All-Male Board Size	0.019***	0.141***	0.009*	0.059+	0.021***	0.171***
	(0.004)	(0.027)	(0.004)	(0.035)	(0.004)	(0.032)
Control variables?	No	No	Yes	Yes	Yes	Yes
Omitted correlated control variables?			No	No	Yes	Yes
Observations	3,593	3,593	3,593	3,593	3,593	3,593
R^2 (OLS)/McFadden's pseudo-R^2 (logit)	0.008	0.009	0.111	0.132	0.079	0.093

Notes. This table shows results from ordinary least squares regressions and logistic regressions predicting whether all-male boards in a given year added at least one woman to their board in the following year. Standard errors, clustered at the firm level, are in parentheses. When controls are present, the regressions include a firm's market capitalization in a given year, a firm's return-on-assets in a given year, a firm's logged total assets in a given year, a firm's one-year total-shareholder returns in a given year, a firm's market-to-book ratio in a given year, a firm's Tobin's Q in a given year, a firm's number of employees in a given year, a firm's level of institutional ownership in a given year measured using percentage of shares owned by institutions, the logged compensation of a firm's CEO in a given year, the proportion of a firm's board members who are independent in a given year, a firm's board's classification (i.e., whether its members are elected to terms that expire in different years), whether a firm's board has a majority voting system, a firm's board's age diversity as measured by the standard deviation of all members' ages in a given year, a firm's board's tenure diversity as measured by the standard deviation of all members' tenure on the board in a given year, fixed effects for calendar year, and fixed effects for industry that a firm belongs to. When omitting correlated control variables, we excluded the following: firm's market capitalization, firm's logged total assets, firm's Tobin's Q, firm's number of employees, firm's CEO's logged compensation, an indicator for firms in the Information Technology industry, and an indicator for firms in the Financials industry.
$^{+}p < 0.10$; $^{*}p < 0.05$; $^{***}p < 0.001$.

absolute correlation of 0.1 or greater with our independent variable of interest—all-male board size—to avoid multicollinearity issues and the problem that these variables (e.g., number of employees) might be proxies for board size rather than appropriate controls.[17] As shown in Table 1, Model (5), when we omit these highly correlated control variables, our predictor for board size is still positive and significant ($b = 0.021$, $p < 0.001$), corresponding to about a 2.1 percentage point increase in the likelihood of an all-male board diversifying for

Figure 2. (Color online) From 2007 to 2018, Larger All-Male Boards Were More Likely Than Smaller All-Male Boards to Add Women (But Not Men) to Their Ranks in a Given Year



Notes. Dots represent the proportions of all-male board-year observations for which at least one new female director or new male director was added in the following year, across board sizes, and are scaled based on relative sample size. The lines represent fitted linear regression lines predicting the likelihood of hiring a female member or male member based on the size of the all-male board.

each additional member. As Table 1, Model (6) shows, this result is consistent when we use a logistic regression ($b = 0.171$, $p < 0.001$, $OR = 1.19$).

Are Larger All-Male Boards also More Likely to Add Men? Our analysis has an important limitation: to add a woman to a board, that board must first decide to change its membership. If larger boards change membership more frequently (because of staggered term limits or because the likelihood that at least one board member will leave in a given year increases as the number of members increases), then larger boards will artificially have more opportunities to add women, which could provide an uninteresting explanation for our findings.

To address these concerns, we reran our analysis, examining whether all-male boards are more likely to add *men* to their ranks as their size increases. If our results are simply driven by the fact that larger boards change membership more frequently, then we should expect larger all-male boards to be more likely to add women *and* men to their ranks as their size increases. However, we find that this is not the case. As shown in Table 2, Model (1), an OLS regression predicting whether a man was added to an all-male board using board size as a predictor shows that the size of an all-male board is an insignificant and directionally negative predictor of adding a new male director ($b = -0.004$, $p = 0.452$; see Figure 2). We find the same pattern when using a logistic regression ($b = -0.017$, $p = 0.453$, $OR = 0.98$; Table 2, Model (2)). Model (3) shows that when we add all of our control variables, the coefficient on board size remains negative and is statistically significant ($b = -0.017$, $p = 0.007$), such that all-male boards are 1.7 percentage points *less* likely to add another man to their ranks for each additional member. Model (4) shows that this pattern is consistent when we run a logistic regression ($b = -0.076$, $p = 0.007$, $OR = 0.93$). Finally, as shown in Model (5), when we omit the control variables that had an absolute correlation of 0.1 or greater with our independent variable of interest, the coefficient on board size remains negative, but is insignificant ($b = -0.008$, $p = 0.175$), and Model (6) shows that we find the same pattern when we run a logistic regression ($b = -0.035$, $p = 0.178$, $OR = 0.97$)

Next, to compare the coefficients on all-male board size across the two types of models (one predicting the addition of a woman to a board and the other predicting the addition of a man), we ran a Z-test designed to enable such comparisons (Clogg et al. 1995, Paternoster et al. 1998). First, comparing the two models that did not include control variables, we find that the coefficient for all-male board size differs significantly between the two models ($Z = 3.635$, $p < 0.001$): in other words, all-male board size is a significantly greater predictor of adding a woman to a board than of adding another man. Next, comparing the two models that include all of our control variables, we once again find

Table 2. Larger All-Male Boards Are Not More Likely to Add Men to Their Ranks

	Dependent variable: *Man Added to All-Male Board*					
Variable	Model (1): OLS regression	Model (2): Logistic regression	Model (3): OLS regression	Model (4): Logistic regression	Model (5): OLS regression	Model (6): Logistic regression
All-Male Board Size	−0.004	−0.017	−0.017**	−0.076**	−0.008	−0.035
	(0.005)	(0.023)	(0.006)	(0.028)	(0.006)	(0.026)
Control variables?	No	No	Yes	Yes	Yes	Yes
Omitted correlated control variables?			No	No	Yes	Yes
Observations	3,593	3,593	3,593	3,593	3,593	3,593
R^2 (OLS)/McFadden's pseudo-R^2 (logit)	0.000	0.000	0.020	0.026	0.016	0.020

Notes. This table shows results from ordinary least squares regressions predicting whether all-male boards in a given year added at least one man to their board in the following year. Standard errors, clustered at the firm level, are in parentheses. When controls are present, the regressions include a firm's market capitalization in a given year, a firm's return-on-assets in a given year, a firm's logged total assets in a given year, a firm's one-year total-shareholder-returns in a given year, a firm's market-to-book ratio in a given year, a firm's Tobin's Q in a given year, a firm's number of employees in a given year, a firm's level of institutional ownership in a given year measured using percentage of shares owned by institutions, the logged compensation of a firm's CEO in a given year, the proportion of a firm's board members who are independent in a given year, a firm's board's classification (i.e., whether its members are elected to terms that expire in different years), whether a firm's board has a majority voting system, a firm's board's age diversity as measured by the standard deviation of all members' ages in a given year, a firm's board's tenure diversity as measured by the standard deviation of all members' tenure on the board in a given year, fixed effects for calendar year, and fixed effects for industry that a firm belongs to. When omitting correlated control variables, we excluded the following: firm's market capitalization, firm's logged total assets, firm's Tobin's Q, firm's number of employees, firm's CEO's logged compensation, an indicator for firms in the Information Technology industry, and an indicator for firms in the Financials industry.

**$p < 0.01$.

that the coefficient for all-male board size differs significantly between the two models ($Z = 3.405$, $p < 0.001$). Finally, comparing the two models that excluded the control variables that are highly correlated with all-male board size, we again find that the coefficient for all-male board size differs significantly between the two models ($Z = 4.147$, $p < 0.001$). Thus, we find that all-male board size positively predicts the addition of women, but not men, to a corporate board, consistent with Hypothesis 2.

Does This Extend to All-White Boards? We reran our analyses examining all-White boards and the addition of non-White board members. We used ISS data from 2011 to 2018 because data on race were missing for approximately half of all directors from 2007 to 2010 (even though ISS data were complete for gender during this period). From 2011 to 2018, director race was missing in <1% of observations, which we were able to fill in manually by searching Google and company websites. See Table S9 in the online supplement for a correlation matrix of all variables used in these analyses.

Of the 4,350 all-White board-year observations, 330 (7.6%) involved boards that added at least one non-White director in the following year. Boards ranged in size from 3 to 17, with a median size of 8. Once again, we winsorized board size at the 0.5th and 99.5th percentiles. The winsorized board size distribution ranged from size 5 to 14. Robustness checks where we do not winsorize the board size distribution are reported in the online supplement (Table S10) and yield similar results to those reported below.

As shown in Table 3, Model (1), all-White board size is a significant and positive predictor of whether a non-White director was added to the board in the following year ($b = 0.009$, $p < 0.001$; see online supplement, Figure S4). Thus, we estimate a 0.9 percentage point increase in the likelihood of an all-White board diversifying for each additional board member. As Table 3, Model (2) shows, this result is consistent when we use a logistic regression ($b = 0.127$, $p < 0.001$, $OR = 1.14$). This suggests that larger all-White boards were more likely to add non-White directors to their ranks than smaller all-White boards, which further supports Hypothesis 2. However, as shown in Table 3, Model (3), when we add our full set of control variables, we find that the coefficient on board size is positive, but it is no longer significant ($b = 0.002$, $p = 0.422$), and Model (4) shows that this pattern is consistent when we run a logistic regression ($b = 0.033$, $p = 0.369$, $OR = 1.03$).

Next, as we did for our gender analyses, as a robustness check, we reran our regression model, but omitted any control variables that had an absolute correlation of 0.1 or greater with our independent variable of interest—all-White board size.[18] As shown in Table 3, Model (5), when we omit these control variables that may simply be alternative proxies for board size, our predictor for all-White board size is, again, a positive

Table 3. Larger All-White Boards Are More Likely to Add Non-White Directors to Their Ranks

	Dependent variable: *Non-White Director Added to All-White Board*					
Variable	Model (1): OLS regression	Model (2): Logistic regression	Model (3): OLS regression	Model (4): Logistic regression	Model (5): OLS regression	Model (6): Logistic regression
All-White Board Size	0.009***	0.127***	0.002	0.033	0.005*	0.095**
	(0.002)	(0.028)	(0.002)	(0.037)	(0.002)	(0.032)
Control variables?	No	No	Yes	Yes	Yes	Yes
Omitted correlated control variables?			No	No	Yes	Yes
Observations	4,350	4,350	4,350	4,350	4,350	4,350
R^2 (OLS)/McFadden's pseudo-R^2 (logit)	0.004	0.008	0.043	0.089	0.035	0.065

Notes. This table shows results from ordinary least squares regressions predicting whether all-White boards in a given year added at least one non-White director to their board in the following year. Standard errors, clustered at the firm level, are in parentheses. When controls are present, the regressions include a firm's market capitalization in a given year, a firm's return-on-assets in a given year, a firm's logged total assets in a given year, a firm's one-year total-shareholder-returns in a given year, a firm's market-to-book ratio in a given year, a firm's Tobin's Q in a given year, a firm's number of employees in a given year, a firm's level of institutional ownership in a given year measured using percentage of shares owned by institutions, the logged compensation of a firm's CEO in a given year, the proportion of a firm's board members who are independent in a given year, a firm's board's classification (i.e., whether its members are elected to terms that expire in different years), whether a firm's board has a majority voting system, a firm's board's age diversity as measured by the standard deviation of all members' ages in a given year, a firm's board's tenure diversity as measured by the standard deviation of all members' tenure on the board in a given year, fixed effects for calendar year, and fixed effects for industry that a firm belongs to. When excluding correlated control variables, we omitted the following: firm's market capitalization, firm's logged total assets, firm's Tobin's Q, firm's number of employees, board's tenure diversity, and an indicator for firms in the Financials industry.

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

Table 4. Larger All-White Boards Are Not More Likely to Add White Directors to Their Ranks

	Dependent variable: *White Director Added to All-White Board*					
Variable	Model (1): OLS regression	Model (2): Logistic regression	Model (3): OLS regression	Model (4): Logistic regression	Model (5): OLS regression	Model (6): Logistic regression
All-White Board Size	0.001	0.006	−0.010^{+}	−0.039^{+}	−0.003	−0.004
	(0.004)	(0.017)	(0.005)	(0.021)	(0.005)	(0.018)
Control Variables?	No	No	Yes	Yes	Yes	Yes
Omitted correlated control variables?			No	No	Yes	Yes
Observations	4,350	4,350	4,350	4,350	4,350	4,350
R^2 (OLS)/McFadden's pseudo-R^2 (logit)	0.000	0.000	0.013	0.017	0.009	0.012

Notes. This table shows results from ordinary least squares regressions predicting whether all-White boards in a given year added at least one White director to their board in the following year. Standard errors, clustered at the firm level, are in parentheses. When controls are present, the regressions include a firm's market capitalization in a given year, a firm's return-on-assets in a given year, a firm's logged total assets in a given year, a firm's one-year total-shareholder returns in a given year, a firm's market-to-book ratio in a given year, a firm's Tobin's Q in a given year, a firm's number of employees in a given year, a firm's level of institutional ownership in a given year measured using percentage of shares owned by institutions, the logged compensation of a firm's CEO in a given year, the proportion of a firm's board members who are independent in a given year, a firm's board's classification (i.e., whether its members are elected to terms that expire in different years), whether a firm's board has a majority voting system, a firm's board's age diversity as measured by the standard deviation of all members' ages in a given year, a firm's board's tenure diversity as measured by the standard deviation of all members' tenure on the board in a given year, fixed effects for calendar year, and fixed effects for industry that a firm belongs to. When excluding correlated control variables, we omitted the following: firm's market capitalization, firm's logged total assets, firm's Tobin's Q, firm's number of employees, board's tenure diversity, and an indicator for firms in the Financials industry.

$^{+}p < 0.10$.

and significant predictor of adding a nonwhite director ($b = 0.005$, $p = 0.029$). Here, we see a 0.5 percentage point increase in the likelihood of an all-White board adding a non-White director is associated with each additional board member. Finally, Model (6) shows that this is consistent when we use a logistic regression ($b = 0.095$, $p = 0.003$, $OR = 1.10$).

Once again, our analysis has the limitation that to add a non-White director, an all-White board must first decide to change its membership, and the rate of change might vary across boards of different sizes. To account for this, we ran additional analyses examining whether all-White boards are more likely to add *White* members to their ranks as their size increases. As shown in Table 4, Model (1), an OLS regression predicting whether a White member was added to an all-White board using board size as a predictor shows that the size of the all-White board is an insignificant predictor ($b = 0.001$, $p = 0.744$; see online supplement, Figure S4), and Model (2) shows that this is also the case when using a logistic regression ($b = 0.006$, $p = 0.744$, $OR = 1.01$). Model (3) shows that the size of the all-White board becomes a negative and marginally significant predictor when we add all of our control variables ($b = -0.010$, $p = 0.055$), such that all-White boards are about 1 percentage point *less* likely to add another White person to their ranks for each additional member. Model (4) shows that this is consistent when we run a logistic regression ($b = -0.039$, $p = 0.056$, $OR = 0.96$). Finally, as shown in Model (5), when we omit the control variables that had an absolute correlation of 0.1 or greater with our independent variable of interest, the coefficient on all-White board size remains negative, but is statistically insignificant ($b = -0.003$, $p = 0.569$), and Model (6) shows that this is also the case when using a logistic regression ($b = -0.004$, $p = 0.819$, $OR = 1.00$)

To compare the coefficients on all-White board size across the two types of models (one predicting the addition of a non-White director to a board and the other predicting the addition of a White director), we first ran a Z-test comparing the two models that did not include controls, and we find that the coefficient for all-White board size differs marginally between the two models ($Z = 1.653$, $p = 0.098$): in other words, all-White board size is a marginally significantly greater predictor of adding a non-White director to a board than of adding another White director. Next, we compare the two models that include all of our control variables, and, here, we find that the coefficient for all-White board size differs significantly between the two models ($Z = 2.075$, $p = 0.038$). Finally, comparing the two models that excluded the control variables that are highly correlated with all-White board size, we do not find a significant difference in the coefficients on all-White board size between the two models ($Z = 1.469$, $p = 0.142$). Thus, we find only limited evidence that the size of an all-White board positively predicts the addition of new, non-White directors over and above the addition of new, White directors.

Robustness Checks. We ran additional analyses to check the robustness of our results, all of which can be found in the online supplement. We reran our analyses where the unit of observation was a newly added director to an all-male or all-White board, rather than an all-male board year or an all-White board year, and we find similar results (Tables S11 and S12). We also ran placebo analyses on a meaningless characteristic of board members (whether their age ended with an arbitrary digit) that should not increase the likelihood of selection as board size increases. As expected, we find that homogeneous board size is not a significant predictor of adding a director with this arbitrary characteristic in our placebo analyses (Table S13). Complete details about our placebo analyses can be found in the online supplement.

Study 4B: The Underrepresentation of Large Homogeneous Boards

In Study 4B, we sought to test our predictions about the downstream consequences of decision makers exerting extra effort to diversify larger homogeneous groups. Specifically, we analyze the composition of S&P 1500 corporate boards in 2018 to test whether larger homogeneous boards (i.e., all-male boards and all-White boards) are more underrepresented relative to chance than smaller homogeneous boards (Hypothesis 4). An unusual scarcity of large homogeneous groups would suggest that people react to the negative signals conveyed by large, homogeneous groups by making exceptional efforts to diversify such groups

Method

Data. Our data again come from ISS. We present results using 2018 data, but our findings are also consistent when analyzing data from 2007 to 2017 (see online supplement, Table S14).

Analysis Strategy. We began by analyzing all-male boards. Because larger boards are mechanically more likely to include at least one woman than smaller boards because they have more seats to fill, we compared the observed distribution of all-male boards across board size with the distribution we would *expect* to see if corporate board directors were chosen from existing directors through a gender-neutral selection process. To do this, we calculated the expected proportion of corporate boards that should be all-male for boards of varying size (ranging from 4 to 23 directors in 2018) based on the null hypothesis that companies select from existing directors and do not consider gender diversity when creating their boards. We then compared the expected proportions of all-male boards of each size to the observed proportions of all-male boards of the same size to determine whether larger boards are disproportionately less likely to be all-male than would be expected by chance.

We calculated the expected proportions of all-male boards using a Monte Carlo simulation method that has been validated in past research (Dezső et al. 2016, Chang et al. 2019). In each simulation, we took the existing universe of 2018 S&P 1500 corporate board members and randomly reassigned all directors to new boards. We took as given the number of boards, the size of each board, and the number of board seats held by each individual director, based on the actual 2018 data. Because directors were randomly assigned to new boards, gender was not a factor in the allocation of board seats. This means the simulation process generated a distribution of men and women on corporate boards that we would expect to see if companies ignored gender when selecting board members and if the only people who were qualified to sit on boards were those who had already held board seats. Moreover, this process accounts for the fact that larger boards are mechanically less likely to be all-male than smaller boards due to the greater number of seats to fill. For each simulation, we then calculated the proportion of all-male boards for each board size (e.g., the proportion of boards with eight directors that consisted only of men).

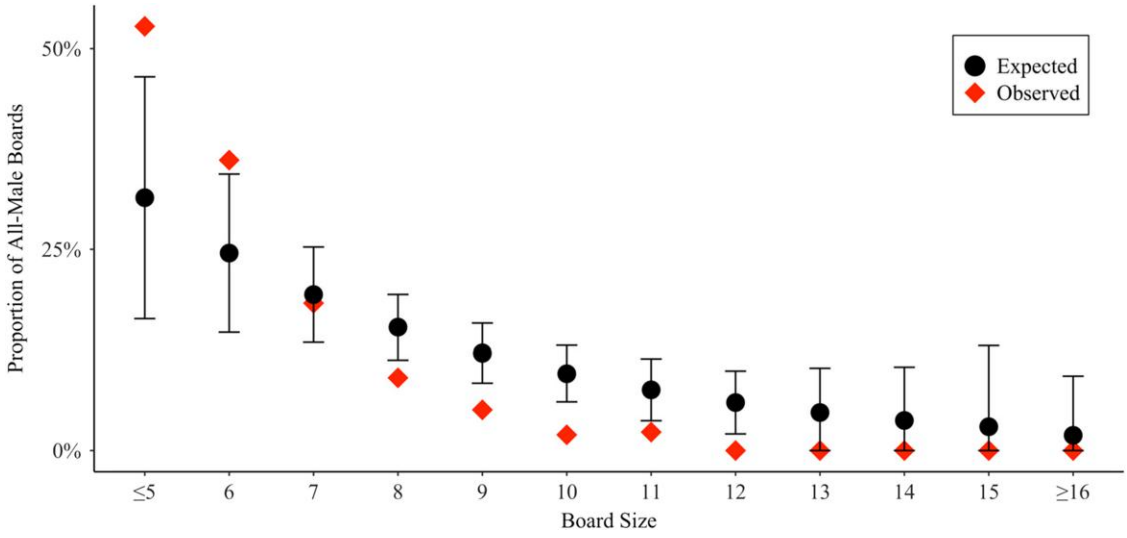
We repeated the simulation process 10,000 times, generating 10,000 possible allocations of directors to corporate boards. For each board size, we calculated the mean proportion of all-male boards across all 10,000 simulations and defined this as the *expected* proportion of all-male boards for a given board size. We also generated 95% confidence intervals around these expected proportions based on our simulated data. Finally, we compared these means to the observed proportions of all-male boards for each board size in 2018.

To test Hypothesis 4, we estimated an ordinary least squares regression to predict the difference between the observed and simulated (i.e., expected) proportions of all-male boards, where our main predictor was board size. Note that because we only performed the simulations with one year of data, there are no repeated observations of firms.

Results

Summary Statistics. The S&P 1500 consisted of 1,506 corporate boards in 2018. Boards in 2018 ranged in size from 4 directors to 23 directors, with an average of 89 boards of each board size, and 8.1% ($n = 123$) of boards were all-male. The median number of directors on a board was nine (see the online supplement, Tables S15 and S16 for summary statistics and the distribution of board sizes in the 2018 data). Once again, we winsorized board size at the 0.5th and 99.5th percentiles. The winsorized board size distribution ranged from size 5

Figure 3. (Color online) Larger All-Male Boards Are Increasingly Underrepresented Relative to Expectations in the S&P 1500 in 2018



Notes. This figure depicts the expected proportions of all-male boards based on Monte Carlo simulations and the observed proportions of all-male boards. Error bars depict 95% confidence intervals.

to 16.[19] Robustness checks where we do not winsorize the board size distribution are reported in the online supplement (Figure S5 and Table S14) and yield similar results to those reported below.

Are All-Male Boards More Underrepresented Among Larger Boards? Figure 3 depicts the expected proportions of all-male boards across board size, as well as the observed proportions of all-male boards in the 2018 data. As shown in Table 5, Model (1), an OLS regression predicting the percent difference between the observed and expected proportions of all-male boards shows that as board size increases, all-male boards become significantly more underrepresented than would be expected by chance ($b = -0.148$, $p < 0.001$; see online supplement, Figure S6). This is consistent with Hypothesis 4.[20] For example, 36% of six-person boards are all-male in 2018, but based on our simulations, we would only expect 25% of them to be all-male ($p = 0.015$), meaning that there are 44% more six-person all-male boards than there would be if board seats were allocated to existing directors in a gender-neutral lottery.[21] On the other hand, only 2% of 11-person boards were all-male in 2018, but based on our simulations, we would expect 8% of them to be all-male ($p = 0.002$), meaning that there are 75% fewer 11-person all-male boards than there would be if board seats were allocated to existing directors in a gender-neutral lottery.

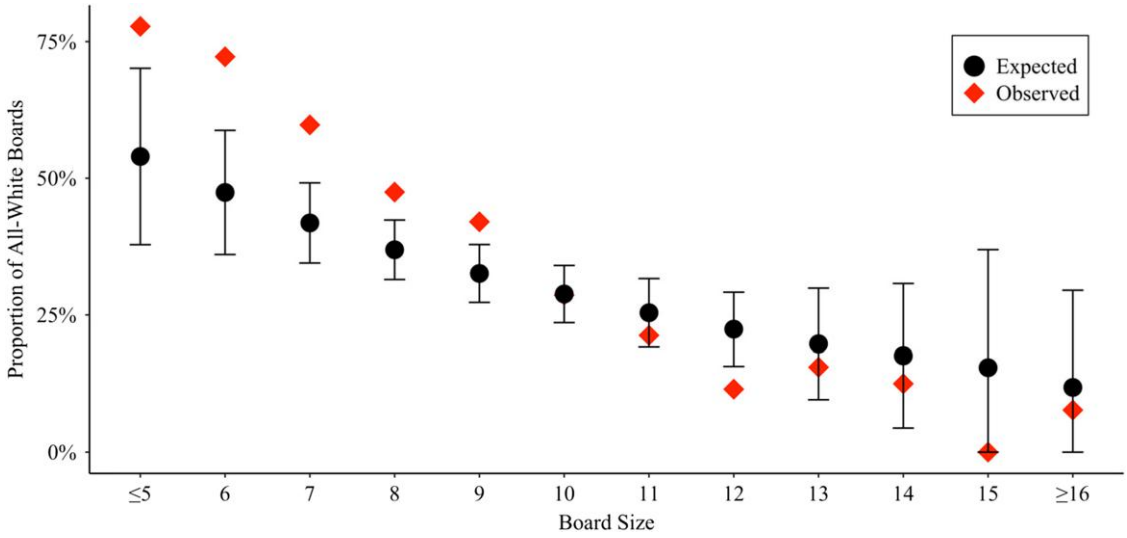
Does This Extend to All-White Boards? We reran our simulations and analyses for all-White boards. In 2018, 37.6% ($n = 567$) boards in the S&P 1500 were all-White. Figure 4 depicts the expected proportions of all-White boards across board size, as well as the observed

Table 5. Levels of Homogeneity Differ Significantly From Expectations as Board Size Increases

	Dependent variable: *Percent Difference Between Observed and Expected Proportions of Homogeneous Boards*	
Variable	Model (1): Gender	Model (2): Race
Board Size	−0.148***	−0.113***
	(0.023)	(0.018)
Constant	1.027**	1.146***
	(0.255)	(0.200)
Observations	12	12
R^2	0.805	0.800

Notes. This table shows regression results from ordinary least squares regressions predicting the percent difference between the observed and expected proportions of all-male boards (Model (1)) and all-White boards (Model (2)) using data from 2018. Expected proportions were computed using Monte Carlo simulations. Standard errors are in parentheses.
**$p < 0.01$; ***$p < 0.001$.

Figure 4. (Color online) Larger All-White Boards Are Increasingly Underrepresented Relative to Expectations in the S&P 1500 in 2018



Notes. This figure depicts the expected proportions of all-White boards based on Monte Carlo simulations and the observed proportions of all-White boards. Error bars depict 95% confidence intervals.

proportions of all-White boards in the 2018 data. As shown in Table 5, Model (2), an OLS regression predicting the percent difference between the observed and expected proportions of all-White boards using 2018 data shows that as board size increases, all-White boards become significantly more underrepresented than would be expected by chance ($b = -0.113$, $p < 0.001$; see online supplement, Figure S7).[22] These results provide further support for Hypothesis 4.[23] For example, 60% of seven-person boards were all-White in 2018, but based on our simulations, we would only expect 42% of them to be all-White ($p < 0.001$), meaning that there were 43% *more* seven-person all-White boards than there would have been if board seats were allocated to existing directors in a race-neutral lottery. On the other hand, only 12% of 12-person boards were all-White in 2018, but, based on our simulations, we would expect 22% of them to be all-White ($p = 0.002$), meaning that there were 45% *fewer* 12-person all-White boards than there would have been if board seats were allocated to existing directors in a race-neutral lottery.

Robustness Checks. To confirm that our results were not an artifact of our simulation strategy, we conducted placebo simulations with a variable that we would not expect to show the same effects (following Chang et al. 2019; details in online supplement). In our placebo simulations, we reran our analysis, but focused on a meaningless characteristic of board members (rather than race or gender): whether their ages ended with an arbitrary digit in 2018 (e.g., whether the board members were 42, 52, 62, etc. in 2018). We defined homogeneous boards as boards that contained no such members. As expected, board size was no longer a significant predictor of the difference between the observed and expected proportions of homogeneous boards on this dimension, suggesting that our results with respect to gender and race are not an artifact of our simulation method or analysis strategy (see online supplement, Figures S9 and S10).

As an additional robustness check, we also reran our simulations within individual industries. Our results were robust across industries for both gender and race (see online supplement, Tables S18 and S19), suggesting that these effects are not driven by particular industries.

Study 4 Discussion

Study 4 extends our experimental findings from Studies 1–3 and tests our theorizing about the importance of group size for diversity-related selection decisions in an important field setting. In Study 4A, we find that larger all-male boards were more likely to diversify by adding female members to their ranks relative to smaller homogeneous boards (Hypothesis 2), but we find only weak evidence of a similar pattern for all-White boards. In Study 4B, we find that among larger boards, all-male and all-White boards were significantly underrepresented relative to chance expectations, and this underrepresentation of homogeneity increases as a function of board size (Hypothesis 4). This means that even in a highly consequential field setting, despite the general prevalence of homogeneous groups in many organizations—and, indeed,

the relative overrepresentation of small homogeneous boards that we document in the S&P 1500—there is a surprising scarcity of large homogeneous corporate boards. This suggests more strategic avoidance of homogeneity by corporate boards as their size grows. In other words, Study 4B suggests that the size of a homogeneous group may be an important enough factor to change substantive selection decisions "in the wild."

General Discussion

Across three experiments and an archival study of U.S. corporate boards from 2007 to 2018, we explore when and why homogeneous groups choose to diversify. In Studies 1 and 2, we manipulate the size of a homogeneous group and demonstrate that a group's size has a causal effect on decision makers' likelihood of diversifying an all-male or all-White group. In Study 3, we test the mechanisms underlying these findings and show that this effect is mediated by perceptions of larger homogeneous groups as (i) more likely to be the product of an unfair selection process; (ii) less diverse; and (iii) more likely to face diversity-related impression management concerns. Finally, in Study 4, we move to the field to establish the external validity and magnitude of these findings. Study 4 shows that for each additional director on a homogeneous board, boards were 1-2 percentage points more likely to diversify their ranks by adding at least one underrepresented member in the year ahead. Moreover, as corporate board size increases, we show that all-male and all-White boards become increasingly underrepresented relative to expectations, suggesting greater strategic avoidance of homogeneity in larger groups.

Theoretical and Practical Implications

Our work makes several contributions to the diversity literature. First, we shed light on when and why decision makers are more or less likely to diversify homogeneous groups. Homogeneous teams are still common in many organizational settings, limiting intergroup contact and exposure to counterstereotypical exemplars. Because such contact and exposure can reduce bias (Lai et al. 2013, Finnegan et al. 2015, Paluck et al. 2019), improving our understanding of the factors that lead homogeneous groups to diversify could be valuable for broader efforts to make organizations more diverse and inclusive. Our findings suggest that decision makers will exert more effort to diversify a group when they worry that it was formed through a biased selection process, judge it as lacking diversity, expect it to face sanctions because it is not diverse enough, or (counter to our theorizing) expect a newly added underrepresented member to wield more influence. These findings suggest that policymakers seeking to increase diversity in organizations may be able to advance their objectives not only by criticizing a lack thereof, but also by calling attention to possible bias in selection processes, making salient the possibility of backlash for lacking diversity or highlighting the positive influence that underrepresented members might wield if hired.

Second, we shed light on the influence of a group's size on hiring decisions and group perceptions. In so doing, we illuminate how a fundamental feature of all groups shapes diversity-related outcomes. At first blush, groups with equivalent levels of diversity might seem indistinguishable from one another, and prior conceptualizations of diversity would treat homogeneous groups of different sizes as indistinct (Harrison and Klein 2007, Meyer 2017). However, we demonstrate that this view is too simplistic: people perceive, judge, and treat homogeneous groups differently depending on their size. In addition, group size influences whether and how aggressively decision makers will exert efforts to diversify homogeneous groups. In the field, we also document a surprising underrepresentation of large homogeneous groups and an overrepresentation of small homogeneous groups on corporate boards, suggesting that group size may influence decision makers' perceptions enough that these high-profile groups exert greater effort to avoid homogeneity the larger they become.

Our work also helps to integrate theory and insights from the judgment and decision-making literature into the diversity literature. We draw upon a core principle of judgment and decision-making research—that people form their beliefs in an intuitively Bayesian manner (Slovic and Lichtenstein 1971, Dawes 1989, Gigerenzer and Hoffrage 1995, Moore and Healy 2008)—and demonstrate that this has important implications for understanding diversity perceptions and forecasting group hiring decisions. We theorize that groups can be viewed as collections of hiring decisions, which means that evaluators receive a stronger signal about the way those decisions are made when evaluating larger groups than smaller groups (Bohnet et al. 2016). Our work suggests that diversity scholars should continue to explore how the size of a group influences diversity-related perceptions and selection decisions, rather than simply holding it constant in experiments or treating it as a control variable in their studies. Moreover, the size of a group is just one potential source of information for evaluators who update their beliefs in a Bayesian manner, and other features of groups that convey relevant information also merit further exploration (e.g., the demographic composition of the members or the distribution of positions of power within the group).

Practically, our theory and findings suggest that decision makers may be more likely to overlook homogeneity when groups are small in size. This may explain the patterns we find in Study 4B, wherein large homogeneous groups are surprisingly scarce (suggesting a

relative overinvestment of effort to diversify those groups), whereas small homogeneous groups are surprisingly abundant (suggesting a relative underinvestment of effort to diversify those groups). Unfortunately, this pattern may have negative consequences. For instance, a startup founded by two White men may not worry about diversifying until the founders realize that all five of their new hires are also White men. However, past research suggests that women and racial minorities may be less willing to join larger homogeneous groups than smaller homogeneous groups (Thomas and Wise 1999, Avery and McKay 2006, Engel et al. 2023). Thus, efforts to diversify homogeneous groups may begin or intensify as they grow in size, but those very groups may be less attractive to women and racial minorities *because* they are large and homogeneous. As a result, groups may miss out on the potential benefits of demographic diversity. For example, positive associations have been shown between diversity and organizational performance, financial performance, and creativity (Richard et al. 2007, Herring 2009, McKay et al. 2009, Homan et al. 2015; cf. Jehn et al. 1997 and Thatcher et al. 2003). Previous work has also shown that majority group members express dissenting ideas more confidently (Phillips and Loyd 2006) and make higher-quality judgments in the presence of underrepresented group members (Sinaceur et al. 2010). In addition, minority group members are more likely than majority group members to share unique information and perspectives in teams (Phillips 2003, Phillips et al. 2004). Homogeneous groups may face disadvantages if they miss out on these potential contributions of demographic diversity.

It is worth noting that although our Bayesian framework is useful for making predictions about diversity-related selection decisions, it is limited in that it does not incorporate the rich social context surrounding diversity in organizations, such as the fact that organizational diversity is often moralized and even politicized (Gasman 2023, Shuman et al. 2023, Winston 2023). These contextual factors may shed additional light on the relative overrepresentation of small homogeneous groups and relative underrepresentation of large homogeneous groups that we observe in Study 4B. It would be valuable for future work to expand upon our Bayesian framework by considering the influence of the social context around diversity on people's judgments and decisions in a given domain.

Relatedly, our findings in Study 3 suggest that impression management concerns may be a driver of choices to diversify homogeneous groups. This may lead organizations to make progress on diversity in terms of numerical representation. However, organizations that focus on numerical representation may not necessarily also take action to create inclusive environments where members of underrepresented groups can thrive—for example, by establishing a positive diversity climate within the organization (Mor Barak et al. 1998, McKay et al. 2007). It would be valuable for future work to explore whether organizations that diversify based on reputational concerns also invest in actually increasing inclusion and belonging.

Extending Our Theory Beyond the Study of Homogeneous Groups

Although our work focuses on homogeneous groups (comprised of members of historically dominant groups in American society), our theorizing can naturally be extended to the study of nonhomogeneous groups. For example, when considering diverse groups (e.g., groups that have achieved gender parity), our theorizing suggests that larger groups still provide stronger signals to evaluators, but now there is more evidence that the group's diversity (rather than its lack thereof) is a meaningful signal and that it effectively represents (rather than diverges from) the underlying population. We would expect this to be reflected in the perceptions formed about diverse groups such that larger diverse groups are more likely to be viewed as (i) the product of a fair selection process; (ii) actually diverse; and (iii) able to avoid diversity-related impression management concerns. Overall, we would predict that decision makers will be less inclined to suggest diversifying larger diverse groups (compared with smaller diverse groups) by adding another underrepresented member.

In a supplemental study (see Study S3 in the online supplement), we conducted an experiment to test this theory. Participants were asked to evaluate two work groups within a hypothetical organization. All participants were shown a group with two members and a group with eight members, and we randomly assigned participants to evaluate either homogeneous all-male groups or gender-diverse groups (with 50% women and 50% men). First, we replicated the patterns from our previous studies in which participants evaluated all-male groups: compared with the two-person homogeneous group, participants judged the eight-person homogeneous group as (i) having a greater need to hire a woman as the next member, (ii) less likely to have employed a fair hiring process, (iii) less diverse, and (iv) more likely to face impression management concerns. However, as we theorized, this pattern reversed when participants were evaluating diverse groups with 50% women. Under these circumstances, compared with the two-person diverse group, participants judged the eight-person diverse group as (i) *less* in need of a female hire, (ii) *more* likely to have a fair hiring process, (iii) *more* diverse, and (iv) *less* likely to face impression management concerns. These supplemental findings illustrate that our theorizing generalizes beyond homogeneous groups. It would be valuable for future work to further expand upon these findings—

for example, by testing how varying group size influences diversity-related perceptions and decisions in groups with some limited diversity (e.g., among groups that are all-male, except for a token female member).

Limitations and Future Directions

We use a multimethod approach to test our theory, combining field data analysis with online experiments to examine our predictions in studies with both strong internal validity (Studies 1–3) and external validity (Study 4). Moreover, each of our studies uses a different decision environment, which helps (i) establish the generalizability of our findings and (ii) confirm they are not an artifact of one particular choice environment. However, an important limitation of our methods is that our experiments only examine decisions made by individuals, whereas our field data reflect corporate board director selection decisions made in a more complex environment with multiple stakeholders weighing in on hires. Moreover, although our experiments test a variety of choice environments and even involve real recommendations in one case (Study 2), they are still low in external validity.

Our field data are also drawn from a single (albeit important) organizational context, so it would be valuable for future research to test whether our findings generalize to other organizational settings. Finally, given that our field data involve many repeated observations from the same firms over time and that the effect sizes we observe are relatively small, we do not have adequate statistical power to test for interactions with our hypothesized mechanisms in our field data. Although our experiments provide support for our hypothesized mechanisms, future work leveraging larger sample sizes to explore mechanism evidence in the field would be valuable.

Our work also focuses primarily on how the size of a group impacts how a group is perceived and who is added next to the group. It would be valuable for future work to examine other organizationally relevant outcomes, such as how the size of a group can influence the *experience* of being a token in an otherwise homogeneous group (Watkins et al. 2019). The experience of joining a homogeneous group and becoming a token or "solo" may be more aversive for women and racial minorities in larger groups, where they may feel particularly isolated and have more difficulty forming friendships (Ozcelik and Barsade 2018). Becoming a token in a larger group might also negatively impact someone's organizational voice (Bowen and Blackmon 2003). Bayesian reasoning and worse experiences in larger, mostly homogeneous groups might also lead women and racial minorities to more strongly believe that organizational decision makers in those larger groups exhibit identity-based biases. Thus, future research could explore how the size of a homogeneous group influences its attractiveness to underrepresented minorities deciding between workplaces (Newburry et al. 2006). Of course, any new member of a group should expect to wield less influence when they are more outnumbered. An interesting open question is whether the reduction in perceived influence is *greater* for members of underrepresented groups, as they face the risks of being tokenized on top of being outnumbered (Kanter 1977, MacCorquodale and Jensen 1993, Watkins et al. 2019).

Another set of questions that future research might explore is whether evaluators' perceptions of diverse groups are accurate and whether this varies by group size and composition. Although the mechanisms that we theorize underlie the effect of group size on diversification decisions are all subjective perceptions, two of the constructs (the perceived fairness of a group's selection process and the perceived amount of influence wielded by a newly added underrepresented member) are more likely to have a "correct" answer than the rest. For example, it is mathematically less likely for a selection process that is unbiased with respect to gender to generate an all-male group as the size of the group increases. Thus, the fairness of a group's selection process is easier to objectively quantify, than, say, the perceived diversity of a group, which can vary across evaluators (Unzueta and Binning 2012, Abascal et al. 2021).

One finding that contradicted our theorizing was that participants in Study 3 were less willing to diversify homogeneous groups when they believed a newly added, underrepresented member would wield less influence. These results suggest that study participants expected homogeneous groups to gain value from diversifying. One explanation for this finding is that people expect homogeneous teams to primarily incur benefits from adding underrepresented members, contrary to our theorizing. Previous work has shown that diversifying groups can have various benefits, such as improving team performance and creativity, increasing members' willingness to express dissenting ideas, boosting the quality of majority group members' judgments, and adding unique information and perspectives to a group (McLeod et al. 1996, Phillips 2003, Phillips et al. 2004, Phillips and Loyd 2006, Richard et al. 2007, Herring 2009, Sinaceur et al. 2010). We theorized that people might see diversity as also having accompanying costs (Mannix and Neale 2005, van Knippenberg and Schippers 2007) and that (biased) decision makers might find it more palatable to add underrepresented group members to teams when they would wield less influence—capitalizing on the impression management benefits of diversity while minimizing its anticipated "costs." However, participants in our study did not conform to our expectations. They appear to have expected homogeneous groups to benefit more from

adding an underrepresented member to their ranks (e.g., adding a woman to an all-male team) if the underrepresented member could exert *more* influence. Future research replicating and expanding upon this pattern would be valuable.

Finally, our work focuses on measuring (Studies 4A and 4B) and manipulating (Studies 1–3) the *actual* size of a group and the consequences of this size for diversification decisions. Because actual group size is generally a given in organizations, to capitalize on the growing appeal of diversification as group size increases, it may be possible to change how large a group "feels" in order to kick-start efforts toward diversification. For example, a homogeneous group could be made to "feel" larger by comparing it to smaller reference groups (Frederick and Mochon 2012, Harris and Speekenbrink 2016). This, in turn, might influence people's beliefs about that group (Schlueter and Scheepers 2010) and lead decision makers to prioritize adding a woman or minority to the group (a hypothesis that would be valuable to test in future research). For example, if there are two homogeneous teams of different sizes in an organization, then using the smaller team as a reference group and making its homogeneity salient to make the other team "feel" bigger might influence decision makers' hiring decisions for the larger team.

Conclusion

Our work sheds light on when and why decision makers are likely to neglect homogeneity in groups versus take action to diversify such groups. We present a Bayesian reasoning framework to explain how a group's size affects perceptions of the group and decisions about whom to hire or promote to join it. In doing so, we illuminate how beliefs about a group's composition and selection processes can drive decisions to diversify. We find consistent evidence across three experiments and a field study that decision makers work harder to diversify larger homogeneous groups. We also document a natural outgrowth of this fact in the context of S&P 1500 corporate boards: larger homogeneous groups are unusually underrepresented relative to expectations, suggesting that organizations work harder to avoid homogeneity in larger groups than smaller ones. By integrating insights from the judgment and decision-making literature and highlighting how a fundamental feature of all groups can influence diversity-related selection decisions, we provide new insights into how organizations can increase their diversity and open new avenues for diversity scholarship.

Acknowledgments

The authors thank Meghan Chung, Joan Dartey, Angelina Pan, and Austin Smith for their valuable research assistance, as well as Kevin Thomas and Mia Wells for library support. The authors also thank Linda Chang, Angela Duckworth, Noah Gans, Sandra Matz, Rebecca Schaumberg, and Maurice Schweitzer for early feedback on this work, as well as the Wharton Behavioral Laboratory for research support. Earlier versions of this work were presented at the 2019 Society for Judgment and Decision Making Conference, the 2020 International Association for Conflict Management Conference, the 2020 Society for Judgment and Decision Making Conference, and the 2021 Academy of Management Annual Meeting, as well as before seminar audiences at American University, Cornell University, Georgetown University, the University of Maryland, and the University of Pennsylvania.

Endnotes

[1] Across all of our experiments, we varied candidate qualifications such that the underrepresented candidate was always objectively more qualified than one of the White male candidates and less qualified than the other White male candidate, and this feature was held constant across all experimental conditions.

[2] We created four sets of three candidates, and participants were randomly assigned to one of these sets. In a pretest (see online supplement Study S2), we find no significant differences in participants' perceptions of how qualified the underrepresented candidate was relative to the most qualified majority candidate in any of the sets.

[3] As a preregistered robustness check, we checked whether our results held if we only analyzed data from participants who remembered the size of the original group shown to them, plus or minus one, and also remembered that the original group contained zero Black people. This did not change our results (in any of our studies).

[4] The online supplement can be found on OSF (which also contains anonymized study data and analysis code): https://osf.io/JZR6D/.

[5] Participants randomly assigned to view larger groups incorrectly recalled seeing more Black people in these (all-White) groups ($b = 0.025$, $p = 0.023$). The direction of this recall error should make our test of Hypothesis 2 more conservative if people are less inclined to diversify groups that they believe already have some diversity.

[6] As a robustness check, we reran our analyses excluding the "group size = 1" condition. When we do this, our results are directionally consistent and become marginally significant (OLS $b = 0.017$, $p = 0.059$; Logistic $b = 0.096$, $p = 0.064$).

[7] We recruited college-educated participants to ensure they were familiar with academic seminars.

[8] We created two sets of three candidates, and participants were randomly assigned to one of these sets.

[9] Participants randomly assigned to view the larger group incorrectly recalled seeing more women in the group ($b = 0.21$, $p < 0.001$). As described in endnote 5, the direction of this recall error should make our test of Hypothesis 2 more conservative.

[10] We created four sets of three candidates, and participants were randomly assigned to one of these sets. In a pretest (see online supplement Study S2), we find no significant differences in participants' perceptions of how qualified the underrepresented candidate was relative to the most qualified majority candidate in all but one of the sets (we find a marginal difference in one set, but our results do not vary when we exclude this set).

[11] Group size was not related to the number of women participants reported seeing in the original (all-male) Innovation Team ($b = -0.026$, $p = 0.532$).

[12] As a robustness check, we reran our analyses excluding the "group size = 1" condition. When we do so, our results remain statistically significant (OLS $b = 0.028$, $p < 0.001$; Logistic $b = 0.148$, $p < 0.001$).

[13] We discovered that 0.5% of directors were labeled with a different gender or race at two different points in time (e.g., the same person was classified as male one year and as female the following year). In all of these cases, we manually fixed the inconsistencies by looking the directors up on Google and company websites.

[14] We selected the specific control variables to include by reviewing all articles published in *Organization Science* since 2015 that used corporate board data and including relevant control variables used in those papers.

[15] Some scholars control for the stock market index to which a firm belongs. However, because a firm's market capitalization is one of the key factors determining stock index membership, we decided to simply control directly for firm market capitalization instead because it provides a continuous measure.

[16] See the online supplement for information on how each of these control variables was accessed.

[17] This led to us omitting the following variables: firm's market capitalization, firm's logged total assets, firm's Tobin's Q, firm's number of employees, firm's CEO's logged compensation, an indicator for firms in the Information Technology industry, and an indicator for firms in the Financials industry.

[18] This led to us omitting the following variables: firm's market capitalization, firm's logged total assets, firm's Tobin's Q, firm's number of employees, board's tenure diversity, and an indicator for firms in the Financials industry.

[19] Note that the size distribution varies from Study 4A because we are using a different set of ISS data (only data from 2018, and all boards instead of only homogeneous boards).

[20] We also computed the absolute percentage point difference (instead of the percent difference) between the observed and expected proportions of all-male boards for each board size and used this difference as an additional dependent measure, producing similar results (see the online supplement, Table S14).

[21] The *p*-values for these comparisons are calculated as the proportion of simulations with a result equally extreme or more extreme than the observed value (e.g., the proportion of 10,000 simulations where at least 36% of six-person boards were all-male; Besag and Clifford 1991) by dividing the number of simulations where an equally or more extreme result was obtained by the total number of simulations run.

[22] Again, when using the absolute difference (rather than the percent difference) between the observed and expected proportions of all-White boards as the dependent measure, we find similar results (see the online supplement, Table S17).

[23] See the online supplement, Table S17 and Figure S8 for results where we do not winsorize the size distribution.

References

Abascal M, Xu J, Baldassarri D (2021) People use both heterogeneity and minority representation to evaluate diversity. *Sci. Adv.* 7(11): eabf2507.

Almeida T, Lordan G (2020) Moving from cheap talk to action: The case of diversity and inclusion. *LSE Business Review* (June 30), https://blogs.lse.ac.uk/businessreview/2020/06/30/moving-from-cheap-talk-to-action-the-case-of-diversity-and-inclusion/.

Apfelbaum EP, Phillips KW, Richeson JA (2014) Rethinking the baseline in diversity research: Should we be explaining the effects of homogeneity? *Perspect. Psych. Sci.* 9(3):235–244.

Avery DR, McKay PF (2006) Target practice: An organizational impression management approach to attracting minority and female job applicants. *Personnel Psych.* 59(1):157–187.

Avery DR, McKay PF, Tonidandel S, Volpone SD, Morris MA (2012) Is there method to the madness? Examining how racioethnic matching influences retail store productivity. *Personnel Psych.* 65(1):167–199.

Bauman CW, Trawalter S, Unzueta MM (2014) Diverse according to whom? Racial group membership and concerns about discrimination shape diversity judgments. *Personality Soc. Psych. Bull.* 40(10):1354–1372.

Beaurain G, Masclet D (2016) Does affirmative action reduce gender discrimination and enhance efficiency? New experimental evidence. *Eur. Econom. Rev.* 90:350–362.

Becker TE (2005) Potential problems in the statistical control of variables in organizational research: A qualitative analysis with recommendations. *Organ. Res. Methods* 8(3):274–289.

Bertrand M, Duflo E (2017) Field experiments on discrimination. Banerjee AV, Duflo E, eds. *Handbook of Economic Field Experiments*, vol. 1 (North-Holland, Amsterdam), 309–393.

Besag J, Clifford P (1991) Sequential Monte Carlo p-values. *Biometrika* 78(2):301–304.

Blader SL (2007) What determines people's fairness judgments? Identification and outcomes influence procedural justice evaluations under uncertainty. *J. Experiment. Soc. Psych.* 43(6):986–994.

Bohnet I, van Geen A, Bazerman M (2016) When performance trumps gender bias: Joint vs. separate evaluation. *Management Sci.* 62(5): 1225–1234.

Bouvy JC, Mujoomdar M (2019) All-male panels and gender diversity of issue panels and plenary sessions at ISPOR Europe. *PharmacoEconom. Open* 3(3):419–422.

Bowen F, Blackmon K (2003) Spirals of silence: The dynamic effects of diversity on organizational voice. *J. Management Stud.* 40(6):1393–1417.

Brown JA, Anderson A, Salas JM, Ward AJ (2017) Do investors care about director tenure? Insights from executive cognition and social capital theories. *Organ. Sci.* 28(3):471–494.

Bunn C (2023) Hamstrung by 'golden handcuffs': Diversity roles disappear 3 years after George Floyd's murder inspired them. *NBC* (February 27), https://www.nbcnews.com/news/nbcblk/diversity-roles-disappear-three-years-george-floyd-protests-inspired-rcna72026.

Carron AV, Spink KS (1995) The group size-cohesion relationship in minimal groups. *Small Group Res.* 26(1):86–105.

Chang EH, Milkman KL, Chugh D, Akinola M (2019) Diversity thresholds: How social norms, visibility, and scrutiny relate to group composition. *Acad. Management J.* 62(1):144–171.

Charness G, Levin D (2005) When optimal choices feel wrong: A laboratory study of Bayesian updating, complexity, and affect. *Amer. Econom. Rev.* 95(4):1300–1309.

Chatman JA, O'Reilly CA (2004) Asymmetric reactions to work group sex diversity among men and women. *Acad. Management J.* 47(2): 193–208.

Chng DHM, Rodgers MS, Shih E, Song XB (2015) Leaders' impression management during organizational decline: The roles of publicity, image concerns, and incentive compensation. *Leadership Quart.* 26(2):270–285.

Clogg CC, Petkova E, Haritou A (1995) Statistical methods for comparing regression coefficients between models. *Amer. J. Sociol.* 100(5):1261–1293.

Colvin C (2022) Once neglected, DEI initiatives now present at all Fortune 100 companies. *HR Dive* (July 20), https://www.hrdive.com/news/2022-fortune-companies-dei/627651/.

Corno L, La Ferrara E, Burns J (2022) Interaction, stereotypes, and performance: Evidence from South Africa. *Amer. Econom. Rev.* 112(12):3848–3875.

Dawes RM (1989) Statistical criteria for establishing a truly false consensus effect. *J. Experiment. Soc. Psych.* 25(1):1–17.

Dezső CL, Ross DG, Uribe J (2016) Is there an implicit quota on women in top management? A large-sample statistical analysis. *Strategic Management J.* 37(1):98–115.

Dunning D, Meyerowitz JA, Holzberg AD (1989) Ambiguity and self-evaluation: The role of idiosyncratic trait definitions in

self-serving assessments of ability. *J. Personality Soc. Psych.* 57(6):1082–1090.

Dutton JE, Dukerich JM, Harquail CV (1994) Organizational images and member identification. *Admin. Sci. Quart.* 39(2):239–263.

Elsbach KD, Sutton RI (1992) Acquiring organizational legitimacy through illegitimate actions: A marriage of institutional and impression management theories. *Acad. Management J.* 35(4):699–738.

Elsbach KD, Sutton RI, Principe KE (1998) Averting expected challenges through anticipatory impression management: A study of hospital billing. *Organ. Sci.* 9(1):68–86.

Else H (2019) How to banish manels and manferences from scientific meetings. *Nature* 573(7773):184–187.

Ely RJ, Thomas DA (2001) Cultural diversity at work: The effects of diversity perspectives on work group processes and outcomes. *Admin. Sci. Quart.* 46(2):229–273.

Engel Y, Lewis T, Cardon MS, Hentschel T (2023) Signaling diversity debt: Startup gender composition and the gender gap in joiners' interest. *Acad. Management J.* 66(5):1469–1500.

Fehr E, Fischbacher U (2004) Third-party punishment and social norms. *Evolution Human Behav.* 25(2):63–87.

Fehr E, Schmidt KM (1999) A theory of fairness, competition, and cooperation. *Quart. J. Econom.* 114(3):817–868.

Fernandez-Mateo I, Fernandez RM (2016) Bending the pipeline? Executive search and gender inequality in hiring for top management jobs. *Management Sci.* 62(12):3636–3655.

Field E, Krivkovich A, Kügele S, Robinson N, Yee L (2023) Women in the workplace 2023. *McKinsey* (October 5), https://www.mckinsey.com/featured-insights/diversity-and-inclusion/women-in-the-workplace.

Finnegan E, Oakhill J, Garnham A (2015) Counter-stereotypical pictures as a strategy for overcoming spontaneous gender stereotypes. *Frontiers Psych.* 6:1291.

Folger R, Bies RJ (1989) Managerial responsibilities and procedural justice. *Employee Responsibilities Rights J.* 2:79–90.

Frederick SW, Mochon D (2012) A scale distortion theory of anchoring. *J. Experiment. Psych. Gen.* 141(1):124–133.

Gallese V, Goldman A (1998) Mirror neurons and the simulation theory of mind-reading. *Trends Cognitive Sci.* 2(12):493–501.

Gasman M (2023) What's with the objection to DEI work? *Forbes* (June 20), https://www.forbes.com/sites/marybethgasman/2023/06/20/whats-with-the-objection-to-dei-work/?sh=393d92e36d35.

Georgeac OAM, Rattan A (2022) Stop making the business case for diversity. *Harvard Business Review* (June 15), https://hbr.org/2022/06/stop-making-the-business-case-for-diversity.

Gigerenzer G, Hoffrage U (1995) How to improve Bayesian reasoning without instruction: Frequency formats. *Psych. Rev.* 102(4):684–704.

Gilliland SW (1993) The perceived fairness of selection systems: An organizational justice perspective. *Acad. Management Rev.* 18(4):694–734.

Gomila R (2021) Logistic or linear? Estimating causal effects of experimental treatments on binary outcomes using regression analysis. *J. Experiment. Psych. General* 150(4):700.

Graffin SD, Carpenter MA, Boivie S (2011) What's all that (strategic) noise? Anticipatory impression management in CEO succession. *Strategic Management J.* 32(7):748–770.

Green J (2020) Goldman to refuse IPOs if all directors are White, straight men. *Bloomberg* (January 23), https://www.bloomberg.com/news/articles/2020-01-24/goldman-rule-adds-to-death-knell-of-the-all-white-male-board.

Green E, Daniels D (2021) Bayesian instinct. Preprint, submitted July 29, https://dx.doi.org/10.2139/ssrn.2916929.

Grieco D, Hogarth RM (2009) Overconfidence in absolute and relative performance: The regression hypothesis and Bayesian updating. *J. Econom. Psych.* 30(5):756–771.

Griffiths TL, Tenenbaum JB (2006) Optimal predictions in everyday cognition. *Psych. Sci.* 17(9):767–773.

Harmeling S (2023) Virtue signaling on race relations only hurts the cause. *Forbes* (January 16), https://www.forbes.com/sites/susanharmeling/2023/01/16/virtue-signaling-on-race-relations-only-hurts-the-cause/?sh=3d329aa452f7.

Harris AJ, Speekenbrink M (2016) Semantic cross-scale numerical anchoring. *Judgment Decision Making* 11(6):572–581.

Harris MM, Lievens F, Van Hoye G (2004) "I think they discriminated against me": Using prototype theory and organizational justice theory for understanding perceived discrimination in selection and promotion situations. *Internat. J. Selection Assessment* 12(1–2):54–65.

Harrison DA, Klein KJ (2007) What's the difference? Diversity constructs as separation, variety, or disparity in organizations. *Acad. Management Rev.* 32(4):1199–1228.

Hegtvedt KA, Johnson C, Ganem NM, Waldron KW, Brody LM (2009) When will the unaffected seek justice for others? Perceptions of and responses to another's injustice. *Australian J. Psych.* 61(1):22–31.

Heilman ME (1983) Sex bias in work settings: The lack of fit model. *Res. Organ. Behav.* 5:269–298.

Hentschel T, Shemla M, Wegge J, Kearney E (2013) Perceived diversity and team functioning: The role of diversity beliefs and affect. *Small Group Res.* 44(1):33–61.

Herring C (2009) Does diversity pay?: Race, gender, and the business case for diversity. *Amer. Sociol. Rev.* 74(2):208–224.

Highhouse S, Brooks ME, Gregarus G (2009) An organizational impression management perspective on the formation of corporate reputations. *J. Management* 35(6):1481–1493.

Holt CA, Smith AM (2009) An update on Bayesian updating. *J. Econom. Behav. Organ.* 69(2):125–134.

Homan AC, Buengeler C, Eckhoff RA, van Ginkel WP, Voelpel SC (2015) The interplay of diversity training and diversity beliefs on team creativity in nationality diverse teams. *J. Appl. Psych.* 100(5):1456–1467.

Institutional Shareholder Services (2019) ISS director data. Accessed July 21, 2019, wrds.wharton.upenn.edu.

Jehn KA, Chadwick C, Thatcher SM (1997) To agree or not to agree: The effects of value congruence, individual demographic dissimilarity, and conflict on workgroup outcomes. *Internat. J. Conflict Management* 8:287–305.

Kanter RM (1977) Some effects of proportions on group life: Skewed sex ratios and responses to token women. *Amer. J. Sociol.* 82(5):965–990.

Kerr NL (1989) Illusions of efficacy: The effects of group size on perceived efficacy in social dilemmas. *J. Experiment. Soc. Psych.* 25(4):287–313.

Kersten D, Mamassian P, Yuille A (2004) Object perception as Bayesian inference. *Annu. Rev. Psych.* 55:271–304.

Kirkland R, Bohnet I (2017) Focusing on what works for workplace diversity. *McKinsey* (April 7), https://www.mckinsey.com/featured-insights/gender-equality/focusing-on-what-works-for-workplace-diversity.

Koch AJ, D'Mello SD, Sackett PR (2015) A meta-analysis of gender stereotypes and bias in experimental simulations of employment decision making. *J. Appl. Psych.* 100(1):128–161.

Kroeper KM, Williams HE, Murphy MC (2022) Counterfeit diversity: How strategically misrepresenting gender diversity dampens organizations' perceived sincerity and elevates women's identity threat concerns. *J. Personality Soc. Psych.* 122(3):399–426.

Kumar S (2018) New study: Almost 70% of professional event speakers are male. *Bizzabo* (November 1), https://www.bizzabo.com/blog/event-gender-diversity-study.

Kunda Z (1990) The case for motivated reasoning. *Psych. Bull.* 108(3):480.

Lai CK, Hoffman KM, Nosek BA (2013) Reducing implicit prejudice. *Soc. Personality Psych. Compass* 7(5):315–330.

Larcker DF, Tayan B (2020) Diversity in the C-suite: The dismal state of diversity among Fortune 100 senior executives. Rock Center

for Corporate Governance at Stanford University Closer Look Series: Topics, Issues and Controversies in Corporate Governance No. CGRP-82, Stanford University, Stanford, CA.

Latané B, Williams K, Harkins S (1979) Many hands make light the work: The causes and consequences of social loafing. *J. Personality Soc. Psych.* 37(6):822–832.

Lawson MA, Martin AE, Huda I, Matz SC (2022) Hiring women into senior leadership positions is associated with a reduction in gender stereotypes in organizational language. *Proc. Natl. Acad. Sci. USA* 119(9):e2026443119.

Leslie LM, Manchester CF, Dahm PC (2017) Why and when does the gender gap reverse? Diversity goals and the pay premium for high potential women. *Acad. Management J.* 60(2):402–432.

Leslie LM, Flynn E, Foster-Gimbel OA, Manchester CF (2023) Happy talk: Is common diversity rhetoric effective diversity rhetoric? *Acad. Management J.* (ja), amj-2021.

Leventhal GS, Karuza J, Fry WR (1980) Beyond fairness: A theory of allocation preferences. Mikula G, ed. *Justice and Social Interaction* (Springer-Verlag, New York), 167–218.

Lotz S, Okimoto TG, Schlösser T, Fetchenhauer D (2011) Punitive vs. compensatory reactions to injustice: Emotional antecedents to third-party interventions. *J. Experiment. Soc. Psych.* 47(2):477–480.

Lount RB Jr, Sheldon OJ, Rink F, Phillips KW (2015) Biased perceptions of racially diverse teams and their consequences for resource support. *Organ. Sci.* 26(5):1351–1364.

Lyness KS, Heilman ME (2006) When fit is fundamental: Performance evaluations and promotions of upper-level female and male managers. *J. Appl. Psych.* 91(4):777–785.

Ma DS, Correll J, Wittenbrink B (2015) The Chicago Face Database: A free stimulus set of faces and norming data. *Behav. Res. Methods* 47(4):1122–1135.

MacCorquodale P, Jensen G (1993) Women in the law: Partners or tokens? *Gender Soc.* 7(4):582–593.

Mack A (2003) Inattentional blindness: Looking without seeing. *Curr. Directions Psych. Sci.* 12(5):180–184.

Mankiw NG (2020) *Principles of Economics* (Cengage Learning, Boston).

Mannix E, Neale MA (2005) What differences make a difference? The promise and reality of diverse teams in organizations. *Psych. Sci. Public Interest* 6(2):31–55.

Maxwell SE, Kelley K, Rausch JR (2008) Sample size planning for statistical power and accuracy in parameter estimation. *Annu. Rev. Psych.* 59:537–563.

McGreevy P (2018) Gov. Jerry Brown signs bill requiring California corporate boards to include women. *Los Angeles Times* (September 20), https://www.latimes.com/politics/la-pol-ca-governor-women-corporate-boards-20180930-story.html.

McKay PF, Avery DR, Morris MA (2009) A tale of two climates: Diversity climate from subordinates' and managers' perspectives and their role in store unit sales performance. *Personnel Psych.* 62(4):767–791.

McKay PF, Avery DR, Tonidandel S, Morris MA, Hernandez M, Hebl MR (2007) Racial differences in employee retention: Are diversity climate perceptions the key? *Personnel Psych.* 60(1):35–62.

McLeod PL, Lobel SA, Cox TH Jr (1996) Ethnic diversity and creativity in small groups. *Small Group Res.* 27(2):248–264.

Mehra A, Kilduff M, Brass DJ (1998) At the margins: A distinctiveness approach to the social identity and social networks of underrepresented groups. *Acad. Management J.* 41(4):441–452.

Meyer B (2017) Team diversity. Salas E, Rico R, Passmore J, eds. *The Wiley Blackwell Handbook of the Psychology of Team Working and Collaborative Processes* (Wiley Blackwell, Hoboken, NJ), 151–175.

Moore DA, Healy PJ (2008) The trouble with overconfidence. *Psych. Rev.* 115(2):502–517.

Moore DA, Small DA (2008) When it is rational for the majority to believe that they are better than average. Krueger JI, ed. *Rationality and Social Responsibility: Essays in Honor of Robyn Mason Dawes* (Psychology Press, Hove, UK), 141–174.

Mor Barak ME, Cherin DA, Berkman S (1998) Organizational and personal dimensions in diversity climate: Ethnic and gender differences in employee perceptions. *J. Appl. Behav. Sci.* 34(1):82–104.

Newburry W, Gardberg NA, Belkin LY (2006) Organizational attractiveness is in the eye of the beholder: The interaction of demographic characteristics with foreignness. *J. Internat. Bus. Stud.* 37(5):666–686.

Nickerson RS (1998) Confirmation bias: A ubiquitous phenomenon in many guises. *Rev. Gen. Psych.* 2(2):175–220.

Ng ES, Wiesner WH (2007) Are men always picked over women? The effects of employment equity directives on selection decisions. *J. Bus. Ethics* 76(2):177–187.

Ozcelik H, Barsade SG (2018) No employee an island: Workplace loneliness and job performance. *Acad. Management J.* 61(6):2343–2366.

Paluck EL, Green SA, Green DP (2019) The contact hypothesis reevaluated. *Behav. Public Policy* 3(2):129–158.

Paternoster R, Brame R, Mazerolle P, Piquero A (1998) Using the correct statistical test for the equality of regression coefficients. *Criminology* 36(4):859–866.

Patrice J (2018) Paul Weiss press release captures everything broken about biglaw in one image. *Above the Law* (December 11), https://abovethelaw.com/2018/12/paul-weiss-press-release-captures-everything-broken-about-biglaw-in-one-image/?rf=1.

Patterson F, Zibarras LD (2011) Exploring the construct of perceived job discrimination in selection. *Internat. J. Selection Assessment* 19(3):251–257.

Petsko CD, Rosette AS (2023) Are leaders still presumed white by default? Racial bias in leader categorization revisited. *J. Appl. Psych.* 108(2):330–340.

Phillips KW (2003) The effects of categorically based expectations on minority influence: The importance of congruence. *Personality Soc. Psych. Bull.* 29(1):3–13.

Phillips KW, Loyd DL (2006) When surface and deep-level diversity collide: The effects on dissenting group members. *Organ. Behav. Human Decision Processes* 99(2):143–160.

Phillips KW, Mannix EA, Neale MA, Gruenfeld DH (2004) Diverse groups and information sharing: The effects of congruent ties. *J. Experiment. Soc. Psych.* 40(4):497–510.

Preacher KJ, Hayes AF (2008) Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behav. Res. Methods* 40(3):879–891.

Pryor Cashman (2017) Pryor Cashman elects four new partners and elevates three to counsel. Accessed February 15, 2020, https://www.pryorcashman.com/news-and-insights/pryor-cashman-elects-four-new-partners-and-elevates-three-to.html.

Richard OC, Murthi BS, Ismail K (2007) The impact of racial diversity on intermediate and long-term performance: The moderating role of environmental context. *Strategic Management J.* 28(12):1213–1233.

Riordan CM, Shore LM (1997) Demographic diversity and employee attitudes: An empirical examination of relational demography within work units. *J. Appl. Psych.* 82(3):342–358.

Robertson DA, Weiss D (2017) In the eye of the beholder: Can counter-stereotypes change perceptions of older adults' social status? *Psych. Aging* 32(6):531–542.

Schaffer BS, Riordan CM (2013) Relational demography in supervisor-subordinate dyads: An examination of discrimination and exclusionary treatment. *Canadian J. Admin. Sci.* 30(1):3–17.

Scheiber N, Eligon J (2019) Elite law firm's all-White partner class stirs debate on diversity. *New York Times* (January 27), https://www.nytimes.com/2019/01/27/us/paul-weiss-partner-diversity-law-firm.html.

Schlueter E, Scheepers P (2010) The relationship between outgroup size and anti-outgroup attitudes: A theoretical synthesis and empirical test of group threat- and intergroup contact theory. *Soc. Sci. Res.* 39(2):285–295.

Shanton K, Goldman A (2010) Simulation theory. *Wiley Interdisciplinary Rev. Cognitive Sci.* 1(4):527–538.

Shemla M, Meyer B, Greer L, Jehn KA (2016) A review of perceived diversity in teams: Does how members perceive their team's composition affect team processes and outcomes? *J. Organ. Behav.* 37:S89–S106.

Shuman E, Knowles E, Goldenberg A (2023) To overcome resistance to DEI, understand what's driving it. *Harvard Business Review* (March 1), https://hbr.org/2023/03/to-overcome-resistance-to-dei-understand-whats-driving-it.

Simmons C (2018) Paul Weiss vows to 'do better' after partner promotions stir diversity debate. *Law.com* (December 18), https://www.law.com/newyorklawjournal/2018/12/18/paul-weiss-vows-to-do-better-after-partner-promotions-stir-diversity-debate/.

Sinaceur M, Thomas-Hunt MC, Neale MA, O'Neill OA, Haag C (2010) Accuracy and perceived expert status in group decisions: When minority members make majority members more accurate privately. *Personality Soc. Psych. Bull.* 36(3):423–437.

Slovic P, Lichtenstein S (1971) Comparison of Bayesian and regression approaches to the study of information processing in judgment. *Organ. Behav. Human Performance* 6(6):649–744.

Spreitzer GM (1995) Psychological empowerment in the workplace: Dimensions, measurement, and validation. *Acad. Management J.* 38(5):1442–1465.

Stevens P (2020) Companies are making bold promises about greater diversity, but there's a long way to go. *CNBC* (June 11), https://www.cnbc.com/2020/06/11/companies-are-making-bold-promises-about-greater-diversity-theres-a-long-way-to-go.html.

Sweeney PD, McFarlin DB (1997) Process and outcome: Gender differences in the assessment of justice. *J. Organ. Behav.* 18(1):83–98.

S&P Dow Jones Indices (2019) S&P Composite 1500, July 2019. Accessed March 29, 2020, https://us.spindices.com/indices/equity/sp-composite-1500.

Thatcher SM, Jehn KA, Zanutto E (2003) Cracks in diversity research: The effects of diversity faultlines on conflict and performance. *Group Decision Negotiation* 12(3):217–241.

Thomas KM, Wise PG (1999) Organizational attractiveness and individual differences: Are diverse applicants attracted by different factors? *J. Bus. Psych.* 13(3):375–390.

Tinsley CH, Wade JB, Main BG, O'Reilly CA (2017) Gender diversity on US corporate boards: Are we running in place? *ILR Rev.* 70(1):160–189.

Triana MDC, Miller TL, Trzebiatowski TM (2014) The double-edged nature of board gender diversity: Diversity, firm performance, and the power of women directors as predictors of strategic change. *Organ. Sci.* 25(2):609–632.

Unzueta MM, Binning KR (2012) Diversity is in the eye of the beholder: How concern for the in-group affects perceptions of racial diversity. *Personality Soc. Psych. Bull.* 38(1):26–38.

Unzueta MM, Knowles ED, Ho GC (2012) Diversity is what you want it to be: How social-dominance motives affect construals of diversity. *Psych. Sci.* 23(3):303–309.

van Knippenberg D, Schippers MC (2007) Work group diversity. *Annu. Rev. Psych.* 58:515–541.

van Knippenberg D, Haslam SA, Platow MJ (2007) Unity through diversity: Value-in-diversity beliefs, work group diversity, and group identification. *Group Dynam.* 11(3):207–222.

Watkins MB, Simmons A, Umphress E (2019) It's not black and white: Toward a contingency perspective on the consequences of being a token. *Acad. Management Perspect.* 33(3):334–365.

Winston A (2023) Why business leaders must resist the anti-ESG movement. *Harvard Business Review* (April 5), https://hbr.org/2023/04/why-business-leaders-must-resist-the-anti-esg-movement.

Aneesh Rai is assistant professor of Management & Organization at the Robert H. Smith School of Business, University of Maryland. He received his PhD from the University of Pennsylvania. His research is focused on using insights from the judgment and decision-making literature to better understand what drives decisions to diversify organizations, as well as exploring interventions to promote positive behavior change.

Edward H. Chang is assistant professor of Negotiation, Organizations, and Markets at Harvard Business School. He received his PhD from the University of Pennsylvania. He researches evidence-based solutions to increase diversity, equity, and inclusion in organizations.

Erika L. Kirgios is assistant professor of Behavioral Science at the Booth School of Business, University of Chicago. She received her PhD from the University of Pennsylvania. She studies the decision-making biases that underlie inequality and designs interventions to improve outcomes for women and members of racial and ethnic minority groups.

Katherine L. Milkman is the James G. Dinan Professor of Operations, Information, and Decisions at the Wharton School, University of Pennsylvania. She received her PhD from Harvard University. Her research explores how insights from economics and psychology can be harnessed to change consequential behaviors (e.g., saving, exercising) for good.