Reseach Article

How Culturally Prevalent Patterns of Nonverbal Behavior Can Influence Discrimination Against Women Leaders

Psychology of Women Quarterly
1–23

DOI: 10.1177/03616843251318964
journals.sagepub.com/home/pwq
Sage

Sarah Ariel Lamer[1], Haley B. Beck[1,2], Leanne ten Brinke[3], and Gillian Michelle Preston[1]

Abstract

We propose that people learn biases against women leaders through patterns of nonverbal behavior depicted in media. Specifically, we hypothesized that (a) people encounter culturally prevalent patterns of nonverbal behavior that favor men leaders over women leaders and (b) seeing patterns of nonverbal behavior favoring men leaders causes people to prefer working under men than women. An analysis of nonverbal behavior directed by and at leaders in 18 popular TV shows revealed that interactions between women leaders and their subordinates were more negative than those between men leaders and their subordinates. In two experimental studies, participants ($N = 193$: 53% women, 47% men, 78% White, $M_{age} = 19.5$ and $N = 237$: 75% women, 25% men, 77% White, $M_{age} = 18.45$) exposed to this nonverbal bias favoring men (vs. a nonverbal bias favoring women) were more likely to choose to work for a White man than a White woman leader. This work has implications for understanding one mechanism through which gender stereotypes of leadership are transmitted and upheld in social groups. *Additional online materials for this article are available on* PWQ's *website at http://journals.sagepub.com/doi/suppl/10.1177/03616843251318964.*

Keywords

leadership, gender discrimination, nonverbal behavior, socialization, media

Over the past decade, the number of women leaders in the United States has been on the rise. For example, a decade ago, women comprised only 4.8% of Fortune 500 CEOs and 20% of U.S. Senators (Catalyst, 2019; McGregor, 2014). Now, the number of Fortune 500 CEOs has doubled (10.4%), and the number of women U.S. senators has risen to 25% (Center for American Women and Politics, 2024; Hinchliffe, 2024). Nonetheless, these numbers are still far from representative of the population, and the advancements women have made in the realm of leadership are nuanced. For example, in political domains, women candidates are slightly preferred over men candidates (Schwarz & Coppock, 2022; Teele et al., 2018). However, this gendered preference is moderated by several critical factors. For example, the bias is stronger for some groups than others (i.e., democrats and independents than republicans; Schwarz & Coppock, 2022), is tempered or even reversed when the position is high-power (e.g., J. L. Smith et al., 2007), and is dependent on women candidates performing a delicate balance of behavior to avoid backlash (Saha & Weeks, 2022). Across domains, women elicit more negative responses than men when engaging in verbal displays of dominance (Rudman & Glick, 2001; Williams & Tiedens, 2016), and when given the option, people are less interested in hiring a woman than a man for a managerial position (Bosak & Sczesny, 2011; Phelan et al., 2008). Men are also granted a wider range of acceptable behavior, while women face greater scrutiny for deviating from nuanced expectations (Bongiorno et al., 2014). Thus, although more women are occupying leadership roles in the United States than ever before, people's attitudes and behavior toward women in leadership still often cause women to be at a disadvantage.

In the current research, we were interested in how media representations shape the beliefs and behaviors that disadvantage women. Media can be powerful in creating and transmitting beliefs and often contains patterns biased toward certain groups or outcomes. For example, U.S. television features a pattern in which women characters are treated

[1]Department of Psychology, The University of Tennessee, Knoxville, TN, USA
[2]MacLean Center for Clinical Medical Ethics, University of Chicago, IL, USA
[3]University of British Columbia - Okanagan, Kelowna, BC, Canada

Corresponding Author:
Sarah Ariel Lamer, University of Tennessee, Knoxville, Austin Peay Building, Knoxville, TN 37996, USA.
Email: slamer@utk.edu

more positively if they are slim (than average-weight) and exposure to this pattern causes women to want to be thinner themselves (Weisbuch & Ambady, 2009). We hypothesized that media has a similar capacity to influence people's beliefs about women in leadership, which we tested with both content analysis and experimental studies.

In recent years, there has been an increase in the number of women shown in counterstereotypical roles in media, such as CEOs, political leaders, and superheroes (Curtis & Cardo, 2018; Hoewe & Sherrill, 2019; Oppenheimer et al., 2003). However, increased representation does not necessarily lead to *equitable* representation. For example, in the 1990s and 2000s, increased media representation of people of color did not lead to increased representation in all roles, but instead to increased stereotypical representation of people of color as lazy, incompetent, and immoral (Mastro & Greenberg, 2000; Monk-Turner et al., 2010). Similarly, when women are featured in media, they are not typically represented in myriad ways but instead shown as stereotypical and sexualized (Signorielli & Bacue, 1999; Sink & Mastro, 2017; S. L. Smith et al., 2010; S. L. Smith & Granados, 2009). In conducting the current research, we sought to understand how interactions with leaders are portrayed nonverbally, with the expectation that interactions with women leaders would be more negative than those with men leaders. Furthermore, we expected that observing negative interactions with women leaders and positive interactions with men leaders (compared to the reverse) would heighten a preference for working under men.

Learning Bias From Nonverbal Emotion

Nonverbal behavior can transmit important social information to observers, and social environments are replete with this source of information. When people are around others, they witness a constant stream of nonverbal behavior, even when no one is speaking. Nonverbal behavior can convey information such as whom to trust, which groups to favor, and how to behave. People also learn biases against certain groups by observing subtle patterns of nonverbal behavior (Brey & Pauker, 2019; Castelli et al., 2008; Lamer et al., 2022; Weisbuch & Ambady, 2009). Furthermore, nonverbal behavior can carry bias similar to *or* absent from explicit sources. For example, popular television shows in 2009 contained a bias in which Black characters received more negatively valenced nonverbal behavior than White characters (Weisbuch et al., 2009). This bias was not reflected in the verbal script, leading to a subtle mechanism through which anti-Black bias can persist. We hypothesized that individuals could derive gender preferences in leadership by similarly observing patterns of nonverbal behavior present when leaders interact with subordinates.

People learn a great deal by observing patterns in others' nonverbal behavior. For example, White children who watched video clips wherein a White actor expressed positive (vs. negative) nonverbal emotion toward a Black actor later endorsed more positive attitudes about Black people in general, corresponding to the nonverbal emotions that they witnessed (Castelli et al., 2008). Whether children heard positive or neutral verbal content in the video had no significant effect on their attitudes. Along a similar vein, Lamer and colleagues (2022) found a pattern in children's television wherein gender-stereotypical girls and boys were treated more positively than gender-counterstereotypical children. When girls between 5 and 10 were exposed to this pattern, they felt more pressure to be gender stereotypical. Adults are also subject to televised nonverbal patterns. For example, Weisbuch and colleagues (Weisbuch et al., 2009; Weisbuch & Ambady, 2009) examined patterns of nonverbal emotion in television that led people to endorse race and body size stereotypes. After viewing the culturally prevalent patterns in which characters expressed more nonverbal positivity toward White (Weisbuch et al., 2009) and slim (Weisbuch & Ambady, 2009) characters, participants favored White people, and women expressed a stronger desire to be slim. In short, nonverbal behavior is an essential source of information that people can use to learn social norms.

Two of the above studies tested whether concurrent verbal information contained the same biases as nonverbal behavior and whether concurrent verbal cues could account for attitude change (Castelli et al., 2008; Weisbuch & Ambady, 2009). Castelli and colleagues (2008) manipulated the verbal content of the interactions they showed children, observing no effect. Weisbuch and Ambady (2009) measured whether the verbal content of television show interactions matched the nonverbal bias against average-weight women, with no pattern emerging. Therefore, an important question is *why* nonverbal behavior is so influential that it can overshadow verbal information, and why it may deviate from what is being said aloud.

There are several reasons that nonverbal behavior may hold such an important role in transmitting beliefs. First, people see nonverbal behavior anytime one person sees another, making it a readily available source of information to learn from that cannot be switched off as verbal behavior can. Second, the layperson tends to believe nonverbal behavior is a more authentic expression of someone's attitudes than verbal behavior (Bogaard et al., 2016; Delmas et al., 2019). Verbal behavior can be consciously modulated, and although nonverbal behavior can be modulated to some extent, it is more challenging to do so (DePaulo, 1992). Third, humans are adapted to learning subtle associations in their environments, such as those involving emotion or nonverbal behavior. For example, people often see a smiling face when they hear laughter. The human brain adapts to this pattern such that people experience a visual illusion when they see a smiling face in the presence of laughter; it looks happier than it is (Sherman et al., 2012). In the case of nonverbal patterns, people can quickly learn who typically gives and

receives positive responses and who typically gives and receives negative responses (Brey & Pauker, 2019; Skinner et al., 2017).

Televised Nonverbal Bias

Given the prevalence of nonverbal behavior, and its capacity to shape beliefs, we hypothesized that people learn negative biases about women's (vs. men's) leadership quality from observing patterns of nonverbal behavior embedded in their culturally prevalent environments. We aimed to test our hypothesis by collecting a sample of nonverbal interactions between leaders and subordinates that many people see regularly. We selected popular television shows in the United States as the source of these nonverbal interactions. Popular television provides one avenue for analyzing nonverbal interactions that a large group of people view habitually. Recent estimates suggest that people watch approximately 3 hr of television daily (Stoll, 2022) and 90% of U.S. adults watch TV at least once weekly (Stoll, 2021). Assessing nonverbal behavior from popular television is, therefore, a practical way to collect interactions that large swaths of a population see (Weisbuch et al., 2017). Each person encounters unique nonverbal interactions based on their neighborhood, the time of day they most often encounter people, or what specific shows they watch. However, given the shared nature of popular television, a pattern in this medium could help to explain beliefs shared by a large group (e.g., members of a cultural group). Therefore, we assessed how women and men leaders on television interact with others, and whether a nonverbal behavior pattern present in this medium could contribute to bias against women leaders.

Nonverbal Interactions in Popular Media

The final output of a television show is the result of choices made by a large team, including writers, actors, directors, editors, and producers. Nonverbal bias against women leaders may emerge at any point in the chain of decisions that are made about a scene, such as the writers' decisions about what to include in the script, the actors' choices in a particular shot, directors' guidance for actors, or editors' choices about which shot to use in the final cut. For example, an actor may hold a bias against women who violate prescribed gender roles, which is then reflected in the actor's nonverbal behavior toward women leaders. Alternatively, women leaders may be scripted to behave more dominantly to reify their identity as leaders (Koenig et al., 2011; Vial & Napier, 2018), causing interactions with them to appear more negative. Identifying any single source of nonverbal bias embedded in media is beyond the scope of the current work. However, critical to our hypotheses, we argue that widely distributed media is likely to have important downstream consequences for viewers and society at large, making the biases embedded within media an important topic for study. We hypothesized that interactions between women leaders and their subordinates in media would reflect and shape widely held beliefs about gender and leadership via how women leaders are treated relative to men leaders, how women leaders act or are perceived to act relative to men leaders, or both. We explored these possibilities in the current work, focusing on nonverbal behavior given its subtle and pervasive ability to inform perceivers' beliefs.

The Current Studies

In Study 1, we investigated what cultural messages popular televised U.S. media communicate to viewers about women and men leaders. We measured how women and men leaders and subordinates acted and were treated nonverbally to answer this question: *Do culturally prevalent patterns of nonverbal emotion in televised interactions favor men leaders over women leaders?* We hypothesized that a biased pattern exists, in which women leaders act more negatively toward subordinates than men leaders and/or are treated more negatively by subordinates than men leaders. In Study 2, we selected a subset of silent clips from Study 1 to experimentally manipulate the cultural pattern and answer this question: *Does exposure to a nonverbal pattern favoring men (vs. women) leaders cause people to prefer men leaders?* We hypothesized that when people watched a nonverbal pattern wherein hierarchical interactions with men leaders were more positive than those with women leaders (vs. the reverse), they would be more likely to want to work for a man than a woman boss. In Study 3, we tested the replicability of the effects of nonverbal bias on leader choice. Because we were interested specifically in nonverbal behavior, we used silent clips throughout all three studies to isolate the role of nonverbal behavior.

All procedures complied with institutional guidelines and were approved by the appropriate IRBs (#19-05397-XM and #2010-1524). Informed consent was obtained for all human participants. All data, materials, and analysis codes that can be shared publicly are here: https://osf.io/ck2xp. Following the accepted reporting standards (Appelbaum et al., 2018), we include the processes we used to determine our exclusions, sample sizes, manipulations, and measures. Data were analyzed using R, Version 4.2.2 (R Core Team, 2022) and the packages ggplot2, Version 3.3.5 (Wickham, 2016), lme4 (Bates et al., 2015), and lmerTest (Kuznetsova et al., 2017). Studies 1 and 2 were not preregistered. Study 3 was confirmatory; its preregistration is available to view here: https://osf.io/e6j3v.

Study 1

We content analyzed nonverbal interactions featuring women and men who were leaders and subordinates from popular television shows in the United States. We hypothesized that interactions between women leaders and their subordinates would be more negative than those between men leaders and their subordinates, a nonverbal bias that could arise in how leaders behaved toward subordinates, how

subordinates behaved toward leaders, or both. In addition to hierarchical interactions between leaders and their subordinates, we collected two kinds of control interactions that enabled us to test whether a nonverbal bias against women leaders was present outside of leadership contexts and whether it was a bias against women more generally.

For the first control about leadership contexts, we collected non-hierarchical interactions featuring leaders with non-subordinates, such as an equal or someone outside the work hierarchy. Given the literature on women in leadership (Bongiorno et al., 2014; Eagly & Karau, 2002; Rudman, 1998; Williams & Tiedens, 2016), we expected to observe the nonverbal bias against women specifically when they were behaving as leaders (i.e., when they were interacting with subordinates). Outside of that leadership role, women no longer violate gender role expectations, so nonverbal bias should no longer be observed.

For the second control about gender bias against women more generally, we collected interactions featuring subordinates matched on gender and race to the leader from each TV show. These interactions could have included hierarchical interactions with the leader, but they did not have to. Instead, they were a representative sample of all subordinates' interactions, such as hierarchical interactions with other superiors and non-hierarchical interactions with people outside of the work hierarchy. By comparing the treatment of women and men subordinates of the same gender and race as the leaders, we could test for a generalized nonverbal bias against women. We did not expect women subordinates to be treated more negatively than men subordinates as this would have indicated a general negativity bias against women. Instead, we expected the nonverbal bias to be specific to women *leaders*. We compared these clips in the content analysis described below.

Method

Participants

Sample size was determined prior to data analysis. We consulted similar past studies where participants rated nonverbal behavior in television clips, and high interrater consistency was observed with 15–23 judges per condition (Lamer et al., 2022; Weisbuch & Ambady, 2009; Weisbuch et al., 2009). We aimed for the higher end of this range in the two rating conditions and stopped data collection when this number was reached. No participants were excluded from this study.

Participants were 45 adults (78% [35] women, 20% [9] men, 2% [1] nonbinary; $M_{age} = 18.89$, $SD = .96$). Of the sample, the majority (89% [40]) self-identified as White/European American, with 2% [1] Asian, 4% [2] Hispanic, 2% [1] Indigenous, and 2% [1] Multiracial (Hispanic/White). The sample was majority heterosexual (89% [40]), with 7% [3] identifying as bisexual, 2% [1] lesbian/gay, and 2% [1] pansexual.

Measures

Stimuli. We followed the cultural snapshots approach to develop a rule-based method for selecting clips. Our goal was to select a representative sample of clips featuring women and men leaders and subordinates (see Table 1; Weisbuch et al., 2017). First, we set our population of shows as adult network television shows in the United States featuring women and men leaders that were airing at the time this study was conducted (i.e., had an episode air in 2015). We aimed to collect 10 shows for each gender: 20 shows and 40 characters total. This goal was based on past work (i.e., Weisbuch et al., 2009 used 11 shows and collected 22 characters; Lamer et al., 2022 used 12 shows and collected 48 characters total).

To compile a list of shows, we began by scouring currently and recently available shows on major television networks for those that featured women or men leaders and gender- and race-matched subordinates appearing with similar frequency. Specifically, advanced research assistants chose a pair of women (one leader, one subordinate) or men (one leader, one subordinate) from each show matched on perceived age, race, and sexual orientation. We continued until we could not find new shows that met these criteria. This led to 43 shows, which we then narrowed down by (a) confirming the quality of character selection within each show, (b) eliminating any shows that did not feature a modern workplace hierarchy, (c) excluding any shows with extreme violence or gore, and (d) matching shows with a woman leader to a show of the same genre with a man leader. See Supplementary Materials for more detailed information on show selection, including a table of the excluded shows with reasons for exclusion.

Once we narrowed down the shows based on the above rules, we then paired each show featuring a woman leader with a show featuring a man leader. As possible, we ensured that the woman leader was similar in age, race, and sexual orientation to the man leader and that the two shows were of the same genre. For example, we matched *Grey's Anatomy* (leader: Meredith Grey—a White, heterosexual woman in her 30s starring in a medical drama) with *Chicago Med* (leader: Will Halstead—a White, heterosexual man in his 30s starring in a medical drama). This process yielded 18 shows.

Given potential interactions with leader race, we had aimed to select as many popular TV shows as possible with Black leaders within the bounds of our selection criteria. Agency backlash explains that women who act agentically (dominant, independent, and assertive) and not communally (warm and friendly) tend to be evaluated negatively (Rudman, 1998). However, agency backlash effects have been moderated by race and are most applicable when evaluating White women and men (Livingston et al., 2012; Rosette et al., 2016). Although White women experience negative reactions to their agency, Black women tend to

Table 1. Application of the Cultural Snapshots Methodology.

Step		Description	Applied to the Current Work
Part 1: Content analysis	Identify cultural pattern of interest	Identify a pattern that the researcher hypothesizes to be present in shared environments	More positive nonverbal behavior in hierarchical interactions with men than women leaders
	Identify human population	Identify the population exposed to this hypothesized pattern	U.S. Americans
	Identify environment	Identify an environment that is commonly and frequently encountered by the population of interest	Televised media on major networks with a recent or currently airing season (2015/2016) that includes modern workplace contexts
	Identify exemplars	Identify the population of exemplars that define the environment	Popular scripted television programs that feature a woman or man leader in a modern workplace hierarchy along with a gender- and race-matched subordinate
	Identify time and location	Identify a representative sample of times and locations to ensure that snapshots collected of each exemplar can be used to estimate perceived culture	4 10-s silent video samples from each of 3 episodes for each character
	Identify and code variables of interest	The variable(s) of interest are defined by the research question and should be coded with an eye toward potential confounding variables	Nonverbal emotional expressions by and toward women and men leaders and subordinates
Part 2: Experiment	Generate experimental conditions	Select (or edit) snapshots such that there is one conditioning containing snapshots consistent with the culturally prevalent pattern and another set without that pattern or with a reverse pattern	Pro-man leader condition (containing the most positive hierarchical interactions between men leaders and their subordinates and the most negative hierarchical interactions between women leaders and their subordinates) Pro-Woman Leader condition (containing the most positive hierarchical interactions between women leaders and their subordinates and the most negative hierarchical interactions between men leaders and their subordinates)
	Test exposure to pattern on outcome variables of interest	Identify variables of interest and test after exposure to experimental condition	Leadership selection (i.e., who would the participants want to work under, a woman or a man)

Note. Adapted from Weisbuch et al. (2017).

be treated more like White *men* in this particular respect: they are conferred similarly high status when acting in communal as dominant ways. However, given the limited diversity in popular media portrayals of leaders at the time, we could only select four mainstream shows with Black leaders (i.e., *How to Get Away with Murder*, *Blindspot*, *Power*, and *Brooklyn 99*). We included these shows to extend the conversation on representations of Black leaders, recognizing that there would be limitations in testing interactions with race.

To clarify our terminology, we used the term *target* when referring to any of the 36 characters who were selected as leaders or subordinates in the shows, and we used the term *interaction partner* when referring to a character who was interacting with one of the target leaders or subordinates in a clip.

After selecting these 18 shows and 36 characters (see Table 2), we sampled clips from three randomly selected episodes of each show and removed any audio. All clips used in these studies were silent to isolate the role of nonverbal behavior. Within each episode, we selected four 10-s clips featuring the leader and four 10-s clips featuring the subordinate. We used the following protocol to select clips. First, we selected three episodes randomly from the current airing season. Then, we split each episode evenly into quarters. From each quarter, we extracted the first 10-s clip in which a target was clearly visible and interacting with one or more interaction partners (see Figure 1). This process was repeated for each of the three episodes. With four clips per episode and three episodes per character, this yielded 12 clips per character. If there was a section without viable clips of both leader and subordinate, we skipped that section. We replaced it with another section

Table 2. Characters and Shows Sampled in Study 1.

Shows with Women Leaders			Shows with Men Leaders			Matched Characteristics Across both Shows		
Show	Leader	Subordinate	Show	Leader	Subordinate	Target Race	Target Sexual Orientation	Show Type
The Catch	Alice Vaughan	Sophie Novak	Blue Bloods	Frank Reagan	Garrett Moore	White	Straight	Crime Drama
Code Black	Leanne Rorisch	Christa Lorenson	Heartbeat	Jessie Shane	Casey Callahan	White	Straight	Medical Drama
Grey's Anatomy	Meredith Grey	Jo Wilson	Chicago Med	Will Halstead	Connor Rhodes	White	Straight	Medical Drama
Limitless	Naz	Rebecca	Bones	Seeley Booth	James Aubrey	White	Straight	Crime Drama
Madam Secretary	Elizabeth McCord	Nadine Toliver	NCIS	Gibbs	McGee	White	Straight	Crime Drama
Parks and Recreation	Leslie Knope	April Ludgate	Last Man Standing	Mike Baxter	Kyle	White	Straight	Comedy
Veep	Selina Meyer	Amy Brookheimer	Superstore	Glenn	Jonah	White	Straight	Comedy
How to Get Away with Murder	Annalise Keating	Michaela Pratt	Power	Ghost	Shawn	Black	Straight	Crime Drama
Blindspot	Bethany Mayfair	Tasha Zapata	Brooklyn 99	Ray Holt	Terry Jeffords	Black Leader; Latina Subordinate	Lesbian/Gay Leader; Straight Subordinate	Crime Drama/ Comedy

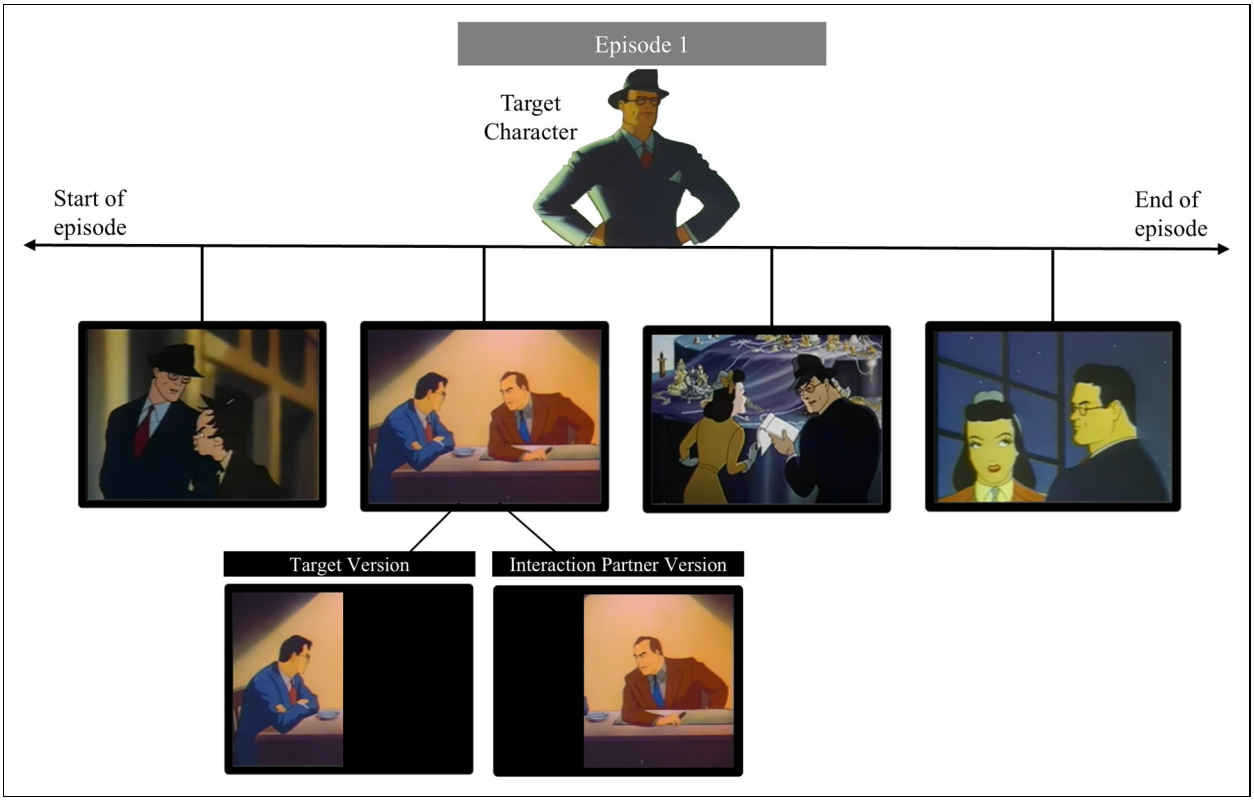
Note. It was not always possible to match the race of a leader to the race of a subordinate from the same show due to a lack of diverse characters and/or lack of adequate screen time for subordinate characters who matched the leader's race. Therefore, one leader/subordinate pair is not race-matched: "Bethany Mayfair" and "Tasha Zapata" from *Blindspot.* Additionally, the leader/subordinate pairs from *Blindspot* and *Brooklyn 99* are not matched within show on sexual orientation for the same reason. Finally, Annalise Keating is bisexual, an identity she verbalized in the show during the last season of How to Get Away With Murder. All clips were drawn from episodes prior to this disclosure, so we coded her as straight to be most consistent with how viewers were likely to categorize her.

from a different episode for both leader and subordinate characters. Note that leaders and subordinates could have been interacting with each other in a single clip. In this case, the clip only counted for one of the targets—whoever appeared first. The other character(s) were considered interaction partners. Following this systematic clip selection protocol, for some characters, we exhausted the episodes from the current airing season before reaching 12 clips (which would have yielded 432 clips). Thus, we collected 414 clips in total.

Interaction Partner Role. To make inferences about how positively subordinates (vs. non-subordinates) treated a leader, we determined which clips featured *hierarchical interactions* whereby leaders interacted with subordinates (i.e., anyone below them in the work hierarchy) and subordinates interacted with leaders (i.e., anyone above them in the work hierarchy) and which clips featured *non-hierarchical interactions* whereby leaders and subordinates interacted with equals or people outside of the work hierarchy. We allowed clips to be taken from any scene (e.g., work context, interaction at home, shopping at the grocery store), and thus, interaction partners could have included anyone with whom the target interacted (e.g., a stranger on the street, a target's coworkers). Note that although we had chosen a leader and a subordinate from each show as target characters, these two did not need to interact for a clip to be considered hierarchical. There were many subordinates for a leader to interact with and many leaders for a subordinate to interact with in a show.

Four advanced undergraduate research assistants coded each interaction partner's role in the clips. Interaction partners in each clip were coded for the number of subordinates, leaders, equals, and people outside of the work hierarchy who were present. For example, in *Grey's Anatomy,* Meredith Grey was the selected woman leader and head of general surgery in the season sampled. She often interacted with Jo Wilson (a subordinate), Arizona Robins (an equal), and Miranda Bailey (a superior). Because accurately coding the interaction partner role sometimes required in-depth knowledge of the show and changes in status across seasons, two research assistants coded each show independently, and disagreements were resolved through discussion between them. Initial agreement ranged from 77% to 98% by show, with an average of 89%. We then simplified the coding to indicate

Figure 1. Method of Selecting and Editing Clips for Each Target Character (i.e., leader or subordinate). For copyright reasons, we used Clark Kent from Fleischer and Fleischer (1941) as an example, since this show is in the public domain. However, we did not use this show in our study as it was well outside of our sampling timeframe. From each episode, we selected four clips of each target character, one from each quarter of the episode, and removed the audio. Each clip featured an interaction with at least one other character (i.e., interaction partner). Then, for each 10-s clip, we made a version with only the target character visible and a version with only the interaction partner(s) visible to be rated by separate sets of participants.



which clips included only hierarchical interactions (i.e., leaders interacting with subordinates) and which did not (e.g., a mix of leaders and subordinates, outsiders, or people in similar positions to the target). We excluded any clips ($n = 53$) in which interaction partners were present that contradicted the target's identity as a leader or subordinate. For example, we excluded clips in which a target character was interacting with both a subordinate and a superior (e.g., Meredith Grey interacting with both Jo Wilson and Miranda Bailey) and clips in which an interaction partner was present who reversed the target's identity (e.g., Meredith Grey, a leader, interacting with only Miranda Bailey, her superior). In these clips, the target character's role was inconsistent or ambiguous. We conducted analyses on the remaining 361 clips.

Rating Strategy. Our goal was to assess how positively other characters (i.e., interaction partners) treated the leader and subordinate in each show and how positively the leaders and subordinates themselves behaved. To do this, participants rated the clips on positivity and liking. If shown the original clips, the identity or emotional expression of the non-focal character(s) could have skewed participants' evaluation of the focal characters' emotions. Therefore, participants evaluated an edited version of each clip in which only targets (i.e., the leader or subordinate) or interaction partners were visible (see Figure 1). We only wanted each participant to see a clip once, so participants rated either clips of targets or interaction partners. Participants were asked to evaluate "How did the visible character(s) behave towards the 'unseen' character?" from 1 (*extremely negatively*) to 7 (*extremely positively*) and "How much did the visible character(s) like or dislike the 'unseen' character?" from 1 (*extremely dislike*) to 7 (*extremely like*). Ratings for these two items were highly correlated ($r = .60$, $p < .001$), so consistent with prior work (Lamer et al., 2022; Weisbuch et al., 2009; Weisbuch & Ambady, 2009), we averaged across the two ratings for each clip that a participant rated, hereafter referring to this variable simply as positivity.

Procedure

Participants were recruited for a study of "Media Psychology" and completed the experiment in person using MediaLab (Jarvis, 2012). They were randomly assigned to code all 414 target or interaction partner clips before completing a demographics questionnaire, being debriefed, and being given partial course credit for their participation. The study took up to two hours.

Results

Positivity

We expected that television contained a nonverbal bias in which subordinates treated women leaders more negatively than men leaders and/or women leaders acted more negatively than men leaders. We collected ratings of target and interaction partner positivity in each clip from multiple participants. Therefore, positivity ratings were nested within clip and rater, rendering cross-classified mixed effects models the best fit for the data (Judd et al., 2012). Mixed models were estimated in R (R Core Team, 2022) with the lme4 package (Bates et al., 2015) using Satterthwaite approximate degrees of freedom (Kuznetsova et al., 2017). Positivity was analyzed as a function of target gender (woman [1] vs. man [−1]), target status (leader [1] vs. subordinate [−1]), rating condition (target [1] vs. interaction partner [−1]), clip type (hierarchical [1] vs. nonhierarchical [−1]), and their interactions. Following recommendations (Barr et al., 2013), we removed slopes that accounted for small amounts of variance (<.002), yielding the following formula: Positivity ~ Target Gender*Target Status*Rating Condition*Clip Type + (Rating Condition|Clip) + (1|Participant). In this model, we included random intercepts of clip and participant in the model and a random slope of rating condition for clip.

We expected that positivity would depend on a target's gender, their status as a leader or subordinate, and whether that leader or subordinate role was salient in the interaction (i.e., their interaction was hierarchical). Providing support for this hypothesis, target gender interacted with target status and clip type, $b = -.09$, $SE = .04$, $t(353.00) = -2.53$, $p = .012$, $\eta^2 = .018$ (see Figure 2).[1] This three-way interaction was not moderated by rating condition, $b = .03$, $SE = .02$, $t(353.00) = 1.21$, $p = .226$, $\eta^2 = .004$, meaning that the same pattern was present when considering how targets acted and how others treated them. Among hierarchical clips, target gender and target status interacted to predict positivity, $b = .11$, $SE = .05$, $t(353.00) = 2.41$, $p = .017$, $\eta^2 = .016$. Women leaders ($M = 3.67$, $SD = 1.09$) were treated more negatively than men leaders ($M = 3.99$, $SD = 1.19$) and acted more negatively ($M = 3.33$, $SD = 1.21$) than men leaders ($M = 3.80$, $SD = 1.19$), $b = -.15$, $SE = .07$, $t(353.00) = -2.26$, $p = .024$, $\eta^2 = .014$. By contrast, women subordinates were neither treated ($M = 4.01$, $SD = 1.14$) nor acted differently ($M = 3.88$, $SD = 0.98$) than men subordinates (interaction partner: $M = 3.95$, $SD = 1.05$; target: $M = 4.15$, $SD = 1.04$), $b = .07$, $SE = .06$, $t(353.00) = 1.11$, $p = .267$, $\eta^2 = .004$. Women and men subordinates were treated similarly regardless of whom they were interacting with, suggesting that there was not a general negative bias against women characters in these shows, but instead a negative bias specific to women leaders in hierarchical interactions. Among non-hierarchical interactions, there was no effect of target gender and status on positivity, $b = -.08$, $SE = .06$, $t(353.00) = -1.36$, $p = .176$, $\eta^2 = .005$. The nonverbal bias was specific to women and men when they were *in* leadership roles. See Table 3 for full regression results.

Robustness Checks

Statistical power is complex and varies substantially based on the expected pattern of the interaction (Giner-Sorolla

Table 3. Regression Results From Study 1.

Term	*b*	*SE*	*df*	*t*	*p*	η^2
(Intercept)	3.93	0.06	129.22	71.25	<.001***	0.016
Target Status	−0.09	0.04	353.00	−2.41	.016*	0.012
Target Gender	−0.08	0.04	353.00	−2.05	.041*	0.009
Clip Type	−0.07	0.04	353.00	−1.84	.066	0.002
Rating Condition	−0.02	0.05	69.39	−0.38	.705	0.000
Target Status × Target Gender	0.01	0.04	353.00	0.37	.714	0.010
Target Status × Clip Type	−0.07	0.04	353.00	−1.85	.065	0.003
Target Gender × Clip Type	−0.04	0.04	353.00	−1.00	.318	0.011
Target Status × Rating Condition	−0.04	0.02	353.00	−1.99	.047*	0.023
Target Gender × Rating Condition	−0.07	0.02	353.00	−2.89	.004**	0.014
Clip Type × Rating Condition	−0.05	0.02	353.00	−2.24	.026*	0.018
Target Status × Target Gender × Clip Type	−0.09	0.04	353.00	−2.53	.012*	0.000
Target Status × Target Gender × Rating Condition	0.01	0.02	353.00	0.26	.794	0.002
Target Status × Clip Type × Rating Condition	−0.02	0.02	353.00	−0.91	0.361	0.000
Target Gender × Clip Type × Rating Condition	−0.01	0.02	353.00	−0.25	.803	0.004
Target Status × Target Gender × Clip Type × Rating Condition	0.03	0.02	353.00	1.21	.226	0.016

Note: Positivity was analyzed as a function of target gender (woman [1] vs. man [−1]), target status (leader [1] vs. subordinate [−1]), rating condition (target [1] vs. interaction partner [−1]), clip type (hierarchical [1] vs. non-hierarchical [−1]), and their interactions, yielding the following formula: Positivity ~ Target Gender*Target Status*Rating Condition*Clip Type + (Rating Condition|Clip) + (1|Participant).
***$p < .001$; **$p < .01$; *$p < .05$.

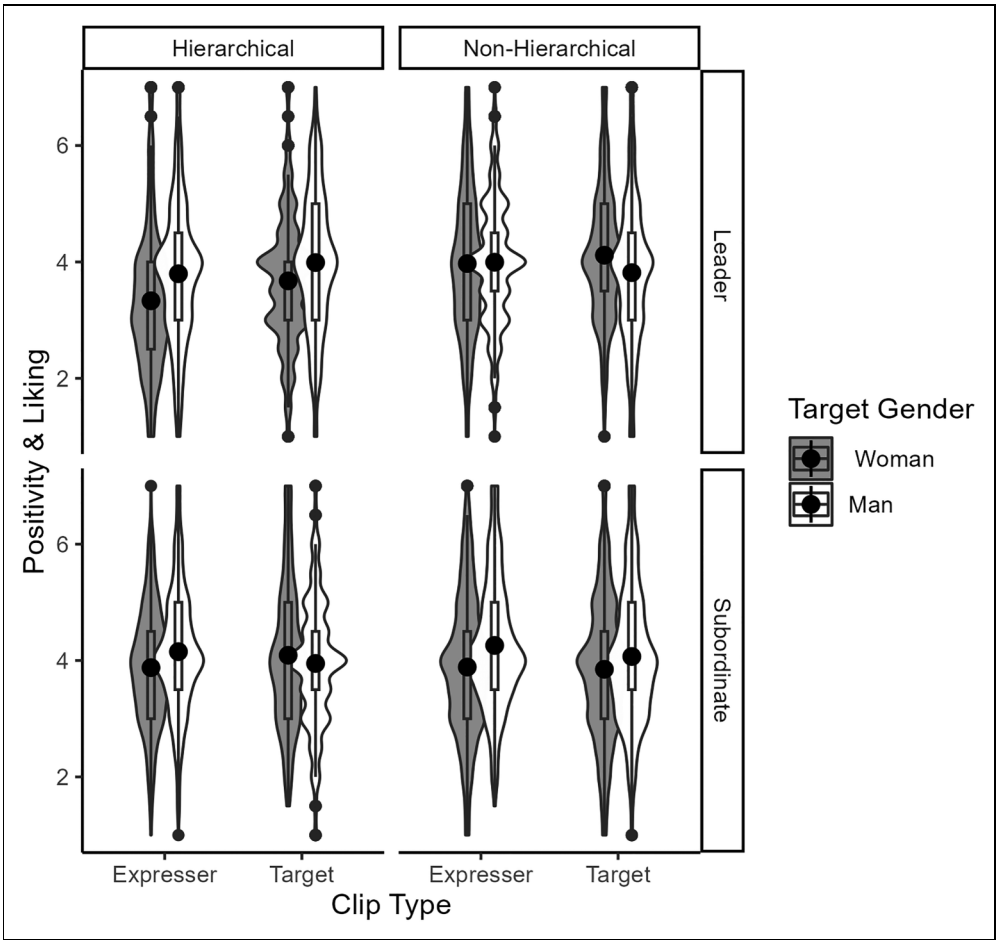
et al., 2024). Therefore, we conducted several kinds of robustness checks to assess the reliability of the reported pattern. First, we conducted sensitivity power analyses for our cross-classified mixed effects model to determine the smallest effect size that could reliably be detected 80% of the time given the current design. Following previous work (Eitan et al., 2018), we used Monte Carlo simulations in the *simr* package (Green & Macleod, 2016). A typical sensitivity analysis provides the smallest effect size that can be detected given the current design. A simulation answers a slightly different question; it provides the power the researchers had to detect an effect of a particular size (Eitan et al., 2018). Therefore, we manually adjusted the effect size in the model until the simulation indicated we had 80% power to reliably detect that effect, holding all other components of our model constant. For the three-way interaction of target gender, target status, and interaction type, we had 80% power to reliably detect an effect of size of $\eta^2 = .020$ in 5,000 simulations (Study 1: $\eta^2 = .018$). In cases like this where sensitivity is very close to the observed effect size, it is important to interpret the model with appropriate caution and run additional robustness checks.

Therefore, we also calculated Type S and Type M error rates (see Gelman & Carlin, 2014). Type S error indicates the likelihood that the sign of the observed effect is the flip of its true effect. Type M error indicates how much the magnitude of the effect reported is likely to be exaggerated relative to the true effect. To do this, the calculations rely on an estimate of the true effect in the population and its standard error drawn from past work alongside the degrees of freedom in the current study. We calculated an estimate of the true effect size by drawing on past work using content analyses of cultural patterns (i.e., Camp et al., 2021; Lamer et al., 2022; Lamer & Weisbuch, 2019; Miyamoto et al., 2006; Weisbuch et al., 2009; Weisbuch & Ambady, 2009), yielding an average *r* of .23 and *SE* of .10. Given these values, it is quite unlikely that the given study would lead to an estimate that is in the wrong direction (0.002%) and is unlikely to be highly exaggerated relative to the true size of the effect (1.26 times). Average effect sizes reported in the literature can overestimate the true size of the effect given publication bias (Gelman & Carlin, 2014). Therefore, we also ran these calculations more conservatively using the smallest effect size among the sample of studies (Lamer & Weisbuch, 2019) in which *r* was .11 (equivalent to $\eta^2 = .013$) and *SE* was .04, yielding a Type S rate of 0.0002% and a Type M rate of 1.14. Altogether, these analyses and error rates suggest that, given the degrees of freedom and the typical size of cultural patterns, the pattern reported in the current work was robust to errors in sign and magnitude. Nonetheless, conservative interpretations of the findings should be made, and further research should be conducted to replicate this effect. Because Type S and M error rates vary when degrees of freedom change, we report Type S and M error rates for each subsequent statistically significant effect that varied in degrees of freedom.

Supplementary Analyses

We also tested the same model above with target race included as a factor, knowing that race intersects meaningfully with leadership evaluations via phenomena like

Figure 2. Nonverbal positivity based on the target character's identity (gender and status), which version of the clip was being rated (target or interaction partner), and who the target was interacting with (hierarchical or nonhierarchical interaction partners). No significant interactions emerged among nonhierarchical clips or among clips featuring subordinates. The key comparisons appear in the upper, left-hand panel of the figure. In hierarchical interactions, women leaders behaved and were treated more negatively than men leaders. Means are shown atop box plots and violin plots to display data distributions in each condition.



agency backlash (Livingston et al., 2012) and the glass cliff (Morgenroth et al., 2020). We hypothesized that nonverbal bias featuring women may be weaker for Black than White leaders. However, given the small number of characters in our sample who were not White ($n = 8$), these analyses were an exploratory test; their full write-up can be found in the Supplementary Materials. In short, the effects reported above remained significant. They were not moderated by target race, suggesting—perhaps surprisingly—that women leaders behaved and were treated more negatively by subordinates than men leaders regardless of race. However, these results should be interpreted with caution given the limited sample and the heterogeneity of sexual orientation among the Black leaders (see Petsko & Bodenhausen, 2019, for an examination of how sexual orientation intersects with race).

We also assessed the gender composition of the interaction partners in each clip. Because we intentionally refrained from limiting the gender of interaction partners in the clips we sampled, it is possible that a gender difference in the subordinates typically shown interacting with women versus men leaders could explain the nonverbal bias in hierarchical interactions. For example, women leaders may act more negatively and be treated more negatively than men leaders because they interact with men more than women (or vice versa). Therefore, we had advanced research assistants code the gender of the interaction partners in each clip. Two research assistants independently counted the number of interaction partners who were women and men in each clip with high interrater reliability (91% agreement); a third research assistant resolved all disagreements with an additional round of independent coding. This coding yielded the proportion of women (versus men) interaction partners in each clip. In hierarchical clips, women leaders were no more likely to interact with women subordinates ($M = .44$, $SD = .47$) than were men leaders ($M = .51$, $SD = .45$), $b = -.03$, $SE = .05$, $t(353) = -0.74$, $p = .460$, partial $\eta^2 = .017$. Therefore, the fact that women leaders acted and were treated more negatively in hierarchical interactions

cannot be explained by a gender imbalance in who they were interacting with.

Altogether, these results suggest that culturally prevalent depictions of leadership contain a negative nonverbal bias involving women. Women leaders and their subordinates are perceived to have more nonverbally negative interactions in popular televised media than men leaders and their subordinates. Of course, behaving negatively does not necessarily mean that leaders were behaving more agentically or less communally. For example, characters could have been rated negatively because they were acting anxious, sad, or fearful. To test the nature of the interactions more precisely regarding relevant leadership qualities, we had an additional set of participants ($N = 48$) evaluate how dominantly or communally the leaders and subordinates behaved in each hierarchical clip from 1 (*not at all*) to 7 (*extremely*). We hypothesized that women leaders would be dominant, but not communal. However, we had no a priori hypotheses about how men leaders would behave since behavioral prescriptions are more flexible for this group (see Bongiorno et al., 2014). Regardless of gender, leaders were perceived to be significantly more dominant ($M = 4.05$, $SD = 1.93$), $b = .45$, $SE = .08$, $t(134.88) = 5.54$, $p < .001$, $\eta^2 = .185$ (Type S = .00003%, Type M = 1.27) and less communal ($M = 3.08$, $SD = 1.60$) than their subordinates (dominance: $M = 3.06$, $SD = 1.72$; communality: $M = 3.76$, $SD = 1.67$), $b = -.31$, $SE = .09$, $t(135.00) = -3.53$, $p < .001$, $\eta^2 = .084$. Interestingly, women leaders were perceived to be significantly *more* dominant ($M = 4.39$, $SD = 1.91$; $b = .34$, $SE = .11$, $t(134.97) = 3.42$, $p < .001$, $\eta^2 = .080$) but no less communal ($M = 2.90$, $SD = 1.56$; $b = -.19$, $SE = .11$, $t(135.00) = -1.76$, $p = .080$, $\eta^2 = .022$) than men leaders (dominance: $M = 3.69$, $SD = 1.89$; communality: $M = 3.28$, $SD = 1.62$). Thus, it appears that leaders' nonverbal behavior in television shows is perceived to be highly dominant and not communal, with women leaders' perceived dominance being greater than that of men leaders' perceived dominance.

Discussion

We found evidence of a subtle cultural pattern of nonverbal bias wherein hierarchical interactions with women leaders were perceived as more negative than those with men leaders. However, women leaders were not always in negative interactions with others. Outside of hierarchical interactions, women and men leaders had similarly valenced interactions, as did women and men subordinates. This suggests that in popular television depictions, women leaders were generally liked as much as men leaders, *except* when they were actually being leaders. Men leaders' behavior and treatment in hierarchical interactions did not significantly differ from the midpoint of the scale, while women leaders' behavior and treatment in hierarchical interactions were significantly below the midpoint. Supplementary analyses indicate that leaders in these shows behaved agentically, but not communally and that women leaders were perceived to be especially dominant in their nonverbal behavior. Extant literature suggests that women are likely subject to negative responses when acting simultaneously low in warmth and high in dominance (Bongiorno et al., 2014; Rudman, 1998; Rudman & Glick, 2001). In hierarchical interactions, women leaders were perceived to act more nonverbally negatively than men leaders and received more negative responses from their subordinates.

The nonverbal pattern biased against women leaders may communicate that women are poor leaders or that interacting with them is unpleasant, making observers of this pattern want to work under men instead. In Study 2, we examined this hypothesis to see whether exposure to this cultural pattern would cause people to prefer men over women as leaders.

Study 2

We conducted Study 2 to determine if seeing more positive hierarchical interactions with men leaders than women leaders could cause people to prefer working under men. Using clip ratings from Study 1, we created two sets of hierarchical clips. As in Study 1, the clips were silent, but we showed the *unedited* version of each clip so that participants could see all the characters. One set of clips contained the prevalent cultural pattern found on television, whereas the other set reversed it. We focused on clips of leaders interacting with subordinates as this is where the role of leader was made salient and where the critical gender difference emerged in nonverbal behavior. Our hypothesis was that participants exposed to a condition that portrayed more negative hierarchical interactions with women leaders compared to men leaders (*Pro-Man Leader Condition*) would be more inclined to choose a man leader to work under compared to those exposed to the opposite pattern (*Pro-Woman Leader Condition*) in which hierarchical interactions with women leaders were more positive than those with men leaders. Consistent with current standards (Chambers et al., 2014), we conducted this initial exploratory experiment to determine the nature, size, and existence of the effect.

Method

Participants

Participants were recruited from introductory psychology courses and participated online for course credit. Past effects of nonverbal bias on behavior and beliefs vary in size, from a Cohen's d of 0.45 (Lamer et al., 2022) and 0.52 (Weisbuch et al., 2009) up to 0.79 (Weisbuch & Ambady, 2009). Therefore, for the current work, we anticipated a medium-sized effect of nonverbal bias ($d = 0.50$ or an Odds Ratio of 2.5; Maher et al., 2013). A power analysis conducted in G*Power to detect this medium-sized effect suggested we needed 208 participants to detect effects on our leadership outcome. Given that this was an initial experimental study, we aimed for this number as a minimum and

collected as many participants as possible in the given academic term. Before exclusions, we met this recruitment goal.

Participants were 193 adults (53% [102] women, 47% [91] men), $M_{age} = 19.5$ ($SD = 2.34$). Of the sample, the majority (78% [150]) self-identified as White/European American, 3% [6] as Asian/Pacific Islander, 12% [23] as Black/African American, 4% [7] as Hispanic/Latine, 1% [2] Middle Eastern, 2% [4] as Multiracial, and 1% [1] as other (no additional information provided). For sexual orientation, 87% [168] identified as heterosexual, 6% [11] as bisexual, 4% [7] as lesbian/gay, 1% [2] as questioning, 1% [1] as demisexual, 1% [2] as asexual, 1% [2] as pansexual, and 1% [1] as other (no additional information provided).[2]

Thirty-nine participants were excluded from this study because they met one or more of the following a priori exclusion criteria: They did not identify as cisgender or transgender women or men ($n = 5$), they started but did not complete the study ($n = 15$; completed less than half of the study), they failed the attention check ($n = 5$), they asked that their data be removed from analyses during the reconsent procedure ($n = 5$), or they failed the manipulation check (i.e., they incorrectly identified half or more of the leaders in the clips; $n = 9$). We considered this last exclusion criterion particularly important because although status is communicated through numerous overt and covert cues (e.g., uniform, visual dominance ratio, expansive posture), the identity of the leader in each clip may not have been clear to all viewers, especially those unfamiliar with a show. Therefore, at the end of the study just before completing the demographics questionnaire, participants rewatched each clip and indicated who they thought the leader in the clip was. They were told to, "describe which character was the leader in that interaction (e.g., you can use their outfit, perceived identity - age, race, gender - or even name if you know it)." These responses were then coded for accuracy, and we excluded anyone who misidentified the leader in at least half of the 16 clips in their condition.

Measures

Stimuli. We used Study 1 clip ratings to create two experimental conditions: a Pro-Man Leader condition and a Pro-Woman Leader condition. We were interested in how watching leaders in their work role influences perceiver behavior, so we selected only from hierarchical clips and instructed participants that the clips they would see included "leaders interacting with people who work for them." For the Pro-Man Leader condition, we started by selecting one clip of each man leader and one clip of each woman leader from our 18 shows. We selected the hierarchical interaction in which each man leader was treated and acted most *positively* and the hierarchical interaction in which each woman leader was treated and acted most *negatively,* limiting those selections to clips that were rated above or below the midpoint, respectively (see Table 4). For the Pro-Woman Leader condition, we did the reverse. We selected the hierarchical interaction in which each man leader was treated and acted most *negatively* and the hierarchical interaction in which each woman leader was treated and acted most *positively*.[3]

There were no eligible hierarchical clips for a few of the 18 shows. For example, in *Last Man Standing*, the leader (Mike Baxter) was often shown interacting with family members, meaning there were no clips in which he was treated negatively only by one of his subordinates. In *Code Black*, none of the clips of the leader (Leanne Rorisch) featured her interacting positively with a subordinate (i.e., all clips were below the neutral midpoint of the scale). Therefore, we dropped these shows from both conditions to ensure each contained clips of the *same* leaders. We supplemented with additional clips from other shows wherever possible to ensure that each condition had enough clips. Altogether, we selected 16 clips for each condition—half featuring women leaders and half featuring men leaders. In each condition, four clips featured Black leaders and 12 featured White leaders. A table of the final clips that were used and their nonverbal positivity ratings are available in the Supplementary Online Materials (Table S2).

Leadership Choice Task. Participants were told we were interested in who they thought would be a good leader, even from minimal information like someone's facial appearance. Participants were to imagine completing a problem-solving task wherein they would work with a team. They were then shown a set of four faces and told that these four faces were team leaders. They were asked to select which leader they would most like to work for. Each set of four faces to choose

Table 4. Average Target and Interaction Partner Positivity for Clips in Each Condition.

	Pro-Woman Leader Condition		Pro-Man Leader Condition	
	Women Leaders ($N = 8$ clips)	Men Leaders ($N = 8$ clips)	Women Leaders ($N = 8$ clips)	Men Leaders ($N = 8$ clips)
Average subordinate positivity toward leader	4.40	3.34	3.26	4.47
Average leader positivity toward subordinate	4.39	3.26	2.95	4.81

Note. Each condition featured 16 clips of leaders interacting with their subordinates. In the Pro-Man Leader Condition, clips featured men leaders acting/being treated positively by subordinates and women leaders acting/being treated negatively by subordinates. In the Pro-Woman Leader Condition, clips featured men leaders acting/being treated negatively by subordinates and women leaders acting/being treated positively by subordinates. The same target characters were featured in both conditions. Positivity was rated on a scale from 1 to 7. Ratings of interaction partner positivity ranged from 1.98 to 6.37. Ratings of target positivity ranged from 1.34 to 6.32.

from contained two women and two men matched on age, trustworthiness, and attractiveness (from the Chicago Face Database norm ratings; Ma et al., 2015). Participants were assigned to view a set of either Black or White faces. To ensure that any effect of condition could not be attributed to idiosyncratic features of just a couple of faces, we selected two sets of White faces and two sets of Black faces matched on the above characteristics. Thus, there were four sets of faces; participants were randomly assigned to view just one set. To explore the effects of condition on Black leader selection relative to White leader selection, we ensured that norming characteristics were also well-matched between the sets. For example, the average age of the eight Black targets was similar to that of the eight White targets. However, since showing multiple sets of faces could elicit order effects or social desirability, participants were assigned to see only one set of faces. Accompanying the faces, they were asked, "Which leader's group would you want to join?"

Procedure

Participants were recruited for a study of "Person Perception" and completed the online experiment using a computer or tablet. They were randomly assigned to view all clips in either the Pro-Man Leader or Pro-Woman Leader condition, having been told that the experimenters were interested in what they thought about people in each clip and rating how negatively/positively the characters felt about each other.[4] They then completed the leader selection task, the manipulation check, and a demographics questionnaire.[5] The study session took approximately 30 min to complete. They were awarded course credit following their participation.

Results

We hypothesized that exposure to the culturally prevalent pattern in which women (vs. men) leaders were treated more negatively would lead people to prefer men (vs. women) leaders, and this preference could be specific to one race. We conducted a binary logistic regression to test our prediction. Specifically, we tested the effects of Clip Condition (Pro-Man Leader [1] or Pro-Woman Leader [−1]), Perceived Leader Race (Black [1] or White [−1]), Participant Gender (Woman [1] or Man [−1]), and their interactions on Leader Choice. We included participant gender in the model because people tend to exhibit gender ingroup biases (Hoyt et al., 2009). Including participant gender as a factor allowed us to account for important variance in leader choice that participant gender could explain.

The main effect of condition was significant and revealed findings consistent with our hypothesis. Participants in the Pro-Man Leader condition were less likely to select a woman leader (44% of the time) than those in the Pro-Woman Leader condition (64% of the time; $b = -.47$, $SE = .16$, $z = -2.99$, $p = .003$, $OR = 0.39$; see Figure 3). There was also a main effect of leader race, which revealed that people selected a woman more often (63% of the time) if the set was White than Black (47% of the time; $b = -.40$, $SE = .16$, $z = -2.58$, $p = .010$, $OR = 0.45$). There was no interaction between clip condition and leader race, $b = -.12$, $SE = .16$, $z = -.78$, $p = .433$, $OR = 0.61$. The effect of clip condition on leader selection was similar regardless of whether participants were assigned to select from among Black or White faces. No other significant effects emerged (see Table 5). Therefore, the nonverbal bias embedded in the clips impacted people's preferences for Black and White men (versus women) leaders similarly.

Robustness Checks

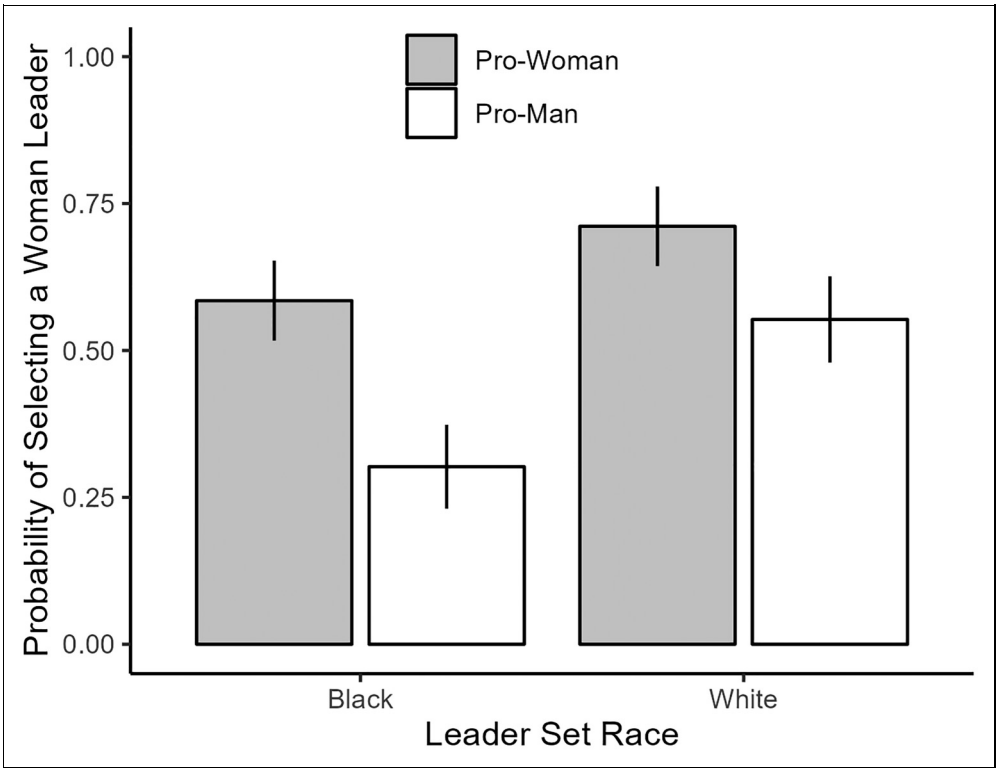
A sensitivity analysis for the effect of condition was conducted using G*Power (Faul et al., 2007) and indicated the minimum detectable effect (with 80% power) given sample size and alpha (i.e., .05). We made the following assumptions drawn from our data: (a) in the Pro-Woman Leader condition, the probability of selecting a woman was .64 (i.e., $\Pr(Y = 1|X = 1)$ $H_0 = .64$), (b) the additional predictors in the model accounted for .036 R^2, (c) and the predictor (condition) was binomial with a balanced design. Under these assumptions, we had 80% power to detect a probability of selecting a woman in the Pro-Man Leader condition of .44 and an OR of 0.43 (here $OR = 0.39$). Thus, our design was sufficiently powered for the observed effect.

Type S and M error calculations were conducted using the average effect size and standard error from studies employing cultural patterns in an experimental design: $d = 0.52$ (equivalent to $OR = 0.40$; Sánchez-Meca et al., 2003), $SE = .24$ (Camp et al., 2021; Lamer et al., 2022; Lamer & Weisbuch, 2019; Miyamoto et al., 2006; Weisbuch et al., 2009; Weisbuch & Ambady, 2009). Given these values, it is quite unlikely that the given studies would lead to an estimate that is in the wrong direction (0.005%) and are unlikely to be greatly exaggerated relative to the true size of the effect (1.32 times). We also ran these calculations more conservatively using the smallest effect size among the sample of studies (Lamer & Weisbuch, 2019) in which d was .20 (equivalent to $OR = 0.70$) and SE was .06, yielding a Type S rate of 0.00002% and a Type M rate of 1.06. These values suggest that the pattern observed in the current work was sufficiently powered and robust to errors in sign and magnitude.

Discussion

We examined how the cultural pattern of nonverbal positivity found in Study 1—in which women leaders were treated and acted more negatively than men leaders—influenced participants' preference for women in positions of leadership. We found evidence that exposure to that pattern led to a greater likelihood of choosing men leaders. Participants were less likely to select a woman leader if choosing from among Black (than

Figure 3. Effect of Clip Condition and Leader in Study 2. There was a main effect of condition such that participants were less likely to select a woman leader in the Pro-Man Leader Condition than the Pro-Woman Leader Condition. Means and standard errors depict predicted values from the fitted model. Error bars depict 1 standard error in either direction of the mean. See supplement for a frequency plot of these data.



White) women and men, but nonverbal bias condition had a similar effect regardless of whether participants were selecting from among Black or White bosses. These data support our theory that exposure to the cultural pattern of bias leads to a stronger preference for men in leadership, affecting people's decisions when selecting leaders. To test the robustness of this pattern of findings in Study 3, we conducted a pre-registered replication (preregistration available here: https://osf.io/e6j3v).

Study 3

Method

Participants

Undergraduate students participated in the study for course credit. The results of a power analysis indicated that a sample of 198 was needed to obtain power of .90. We calculated our power analysis in G*Power using the observed size of the main effect in Study 2. We met this sample size goal before applying our pre-registered exclusion criteria.

Participants were 237 adults (75% [178] women, 25% [59] men), $M_{age} = 18.45$ ($SD = 1.38$). Of the sample, the majority (77% [183]) self-identified as White/European American, with 6% [15] Asian/Pacific Islander, 5% [13] Black, 5% [11] Hispanic/Latine, <1% [1] Middle Eastern, 5% [13] Multiracial, and <1% [1] as other (no other information provided). Regarding sexual orientation, 86% [204] identified as heterosexual, <1% [1] asexual, 10% [23] as bisexual, 2% [5] as lesbian/gay, 2% [4] as pansexual, <1% [1] as queer, and 1% [2] other (no other information provided).

Twenty-two participants were excluded from this study because they met one or more of the following a priori exclusion criteria: They were younger than 18 ($n = 2$), they did not identify as cisgender or transgender women or men ($n = 10$), they did not complete the full study ($n = 2$), they failed the manipulation check (i.e., they incorrectly identified half or more of the leaders in the clips; $n = 2$), or they failed the attention check ($n = 10$). Due to a programming error in PsychoPy, attention check responses did not record for an additional 46 participants. Therefore, given we met our sample size goal, we conducted analyses conservatively by removing data from these participants. However, including them did not meaningfully change the results.

Procedure

Participants watched the clips for their assigned condition (the same Pro-Man Leader and Pro-Woman Leader conditions from Study 2) and completed the same leader selection measure, manipulation check, and demographics from Study 2.[6]

Table 5. Regression Results from Study 2.

Term	*b*	*SE*	*z*	*p*	*OR*
(Intercept)	0.15	0.16	0.99	.321	-
Clip Condition	−0.47	0.16	−2.99	.003**	0.39
Race of Leaders to Choose From	−0.40	0.16	−2.58	.010**	0.45
Participant Gender	0.29	0.16	1.88	.060	1.80
Clip Condition × Race of Leaders to Choose From	−0.12	0.16	−0.78	.433	0.61
Clip Condition × Participant Gender	0.13	0.16	0.83	.405	1.68
Race of Leaders to Choose From × Participant Gender	0.07	0.16	0.48	.632	1.35
Clip Condition × Race of Leaders to Choose From × Participant Gender	0.05	0.16	0.32	.747	1.50

Note: Chosen leader gender (woman [1] vs. man [0]) was analyzed as a function of clip condition (pro-man leader [1] vs. pro-woman leader [−1]), the race of leaders to choose from (Black [1] vs. White [−1]), and participant gender ([1] woman vs. [−1] man) and their interactions, yielding the following formula: Leader Gender ~ Clip Condition*Race of Leaders to Choose From*Participant Gender.

Results

We tested the hypothesis that exposure to the culturally prevalent pattern in which women (vs. men) leaders were treated more negatively would lead people to prefer men (over women) leaders. No main effect of clip condition emerged, $b = -.06$, $SE = .17$, $z = -.36$, $p = .722$, $OR = 0.89$. However, clip condition interacted with leader race, $b = .47$, $SE = .17$, $z = 2.72$, $p = .006$, $OR = 6.46$. Simple effects revealed that the hypothesized effect for White face sets replicated. Participants in the Pro-Man Leader condition were significantly less likely to select a White woman leader than those in the Pro-Woman Leader condition (58% vs. 74% of the time), $b = -.53$, $SE = .26$, $z = -2.03$, $p = .043$, $OR = .35$. However, the effect was not significant if the leaders to select from were Black (see Figure 4). Participants were not significantly more likely to choose a Black woman in the Pro-Woman Leader condition (42% of the time) than in the Pro-Man Leader condition (60% of the time), $b = .41$, $SE = .22$, $z = 1.82$, $p = .069$, $OR = 2.25$. There was a main effect of leader race as there had been in Study 2, which revealed that people were more likely to select a woman if the group was White than if the group was Black (65% vs. 50% of the time; $b = -.37$, $SE = .17$, $z = -2.15$, $p = .032$, $OR = 0.48$). No other main or interactive effects emerged (see Table 6).

Robustness Checks

A sensitivity analysis for the effect of condition was conducted using G*Power (Faul et al., 2007) and indicates the minimum detectable effect (with 80% power) given sample size and alpha (i.e., .05). We made the following assumptions drawn from our data: (a) in the Pro-Woman Leader condition, the probability of selecting a woman was .57 (i.e., $\Pr(Y = 1|X = 1)$ $H_0 = .57$), (b) the additional predictors in the model accounted for .065 R^2, (c) and the predictor (condition) was binomial with a balanced design. Under these assumptions, we had 80% power to detect a probability of selecting a woman in the Pro-Man Leader condition of .38 and an *OR* of 0.47 (here $OR = 0.89$). We also conducted a sensitivity analysis for the simple effect of condition for White face sets. We made the following assumptions drawn from our data: (a) in the Pro-Woman Leader condition, the probability of selecting a White woman was .74 (i.e., $\Pr(Y = 1|X = 1)$ $H_0 = .74$), (b) the additional predictors in the model accounted for .051 R^2, (c) and the predictor (condition) was binomial with a balanced design. Under these assumptions, we had 80% power to detect a probability of selecting a woman in the Pro-Man Leader condition of .56 and an *OR* of 0.45 (here $OR = 0.35$). Thus, the design was sufficiently powered to detect the observed effect of condition on White leaders.

Type S and M error calculations again suggest it is quite unlikely that the given studies would lead to an estimate that is in the wrong direction (0.004%) and are unlikely to be greatly exaggerated relative to the true size of the effect (1.31 times). We also ran these calculations more conservatively using the smallest effect size among the sample of studies (Lamer & Weisbuch, 2019) in which *d* was .20 (equivalent to $OR = 0.70$) and *SE* was .06, yielding a Type S rate of 0.00001% and a Type M rate of 1.05. These values suggest that the pattern observed in the current work was sufficiently powered and robust to errors in sign and magnitude.

Discussion

In Study 3, we observed support for the hypothesis that patterns of nonverbal positivity in brief interactions would shape preferences for working under White women (vs. men). Though participants selected a White woman more than 50% of the time regardless of condition, we nonetheless observed the hypothesized effect of condition. Given the option of selecting a White woman or man, participants exposed to a pattern of nonverbal behavior favoring women (versus men) leaders were significantly more likely to choose a woman than a man leader. We did not see this pattern emerge when participants selected

Figure 4. Effect of Clip Condition and Leader Race in Study 3. People were less likely to select a White woman leader in the Pro-Man Leader than the Pro-Woman Leader condition. The effect of condition on Black leader selection was not significant. Means and standard errors depict predicted values from the fitted model. Error bars depict 1 standard error in either direction of the mean. See Supplementary Materials for a frequency plot of these data.

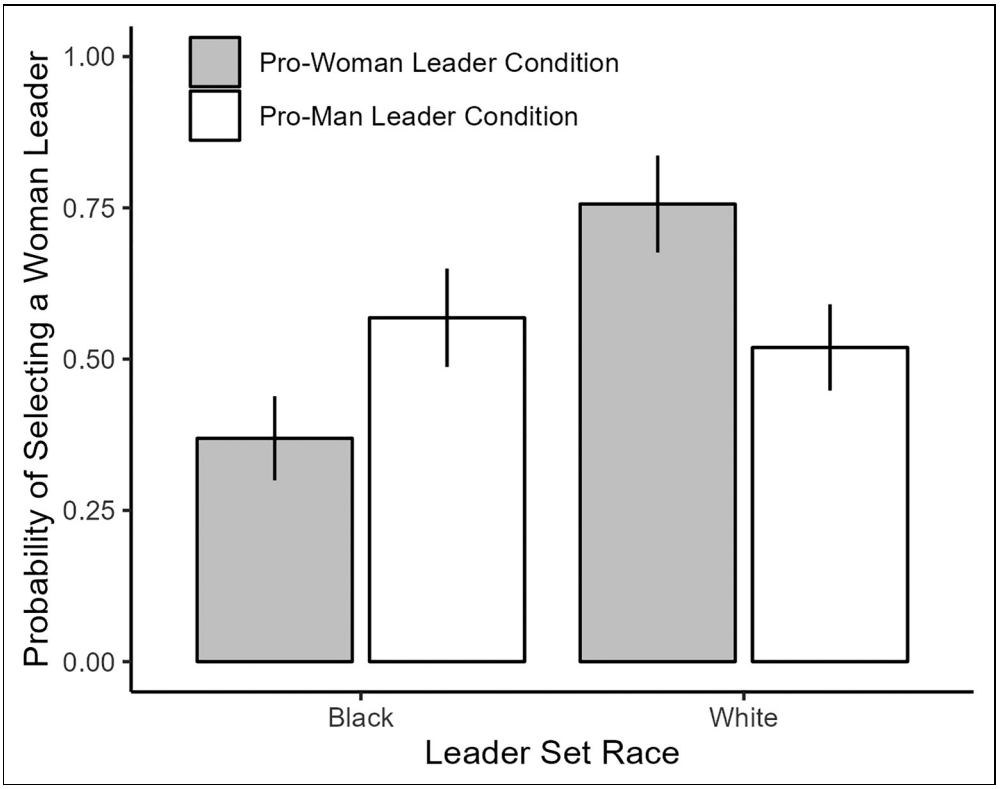


Table 6. Regression Results from Study 3.

Term	*b*	*SE*	*z*	*p*	*OR*
(Intercept)	0.24	0.17	1.38	.166	-
Clip Condition	−0.06	0.17	−0.36	.722	0.89
Race of Leaders to Choose From	−0.37	0.17	−2.15	.032*	0.48
Participant Gender	0.32	0.17	1.84	.065	1.88
Clip Condition × Race of Leaders to Choose From	0.47	0.17	2.72	.006**	6.46
Clip Condition × Participant Gender	0.13	0.17	0.76	.449	1.68
Race of Leaders to Choose From × Participant Gender	0.07	0.17	0.40	.690	1.31
Clip Condition × Race of Leaders to Choose From × Participant Gender	−0.24	0.17	−1.39	.163	0.15

Note: Chosen leader gender (woman [1] vs. man [0]) was analyzed as a function of clip condition (pro-man leader [1] vs. pro-woman leader [−1]), the race of leaders to choose from (Black [1] vs. White [−1]), and participant gender ([1] woman vs. [−1] man) and their interactions, yielding the following formula: Leader Gender ~ Clip Condition*Race of Leaders to Choose From*Participant Gender.

from Black leaders. We discuss how the race of the leaders in the clips or the leaders pictured in the task may have interacted with this effect in the General Discussion below. Overall, these data contribute to the literature on norms regarding gender biases in leadership. The effect of condition on leader selection in White groups emerged above and beyond co-occurring effects, though we discuss these co-occurring effects in more detail below.

General Discussion

When asked, people tend to support the idea of women holding leadership positions (Parker et al., 2018) and sometimes put that idea into action (Schwarz & Coppock, 2022) with many political scientists pointing toward the lack of women running for political positions as reason for their low numbers, not a sustained bias against women being

leaders (Saha & Weeks, 2022). Nonetheless, the data overwhelmingly support the conclusion that women are subject to a more narrow window of acceptable behavior as leaders (Bongiorno et al., 2014; Livingston et al., 2012). We reasoned that patterns of nonverbal behavior in media favor men and convey that men are preferred for leadership roles relative to women. We also hypothesized that media is an important source of information about who to support in leadership roles and that people may learn to hold preferences for men leaders by observing women's and men's interactions when in leadership roles.

We conducted three studies to test these hypotheses. We found evidence that (a) in televised interactions, hierarchical interactions with women leaders are more negative than those with men leaders and (b) participants exposed to this culturally prevalent pattern of nonverbal bias are less likely to want to work for a woman leader than a man leader. These data support our predictions that people can learn preferences for men (over women) leaders through subtle, ubiquitous means and that culturally prevalent, televised representations of women in leadership communicate a preference for having men (versus women) in leadership roles.

A Cultural Pattern of Nonverbal Bias: Implications and Limitations

Study 1's aim was to sample representations that people in the United States encounter most frequently, drawing from the goals of the cultural snapshots approach (Weisbuch et al., 2017). These representations necessarily varied along several other dimensions, such as the leaders' personalities, the show's storyline, or the characters' complexity. Indeed, women are often depicted in systematically different ways than men on television (Sink & Mastro, 2017). These co-occurring cues are embedded in the pattern of interest, and collectively, these portrayals represent some of the most popular televised depictions of leaders. Rather than restricting our character selection so that women and men leaders were more closely matched on all potential covarying cues, we emphasized cultural exposure, selecting the most commonly encountered representations of women and men leaders. When they were aired during the season, viewership for the 18 shows that we sampled ranged from 1.05 million viewers to 15.9 million viewers per show, totaling 114 million viewers. Thus, the nonverbal bias in televised interactions was widely observed.

From our data, we do not know whether gender differences in target positivity (i.e., that women leaders were perceived to be more negative than men leaders in hierarchical interactions) were due to objective differences in leader behavior or due to subjective differences in how participants perceived women's and men's behavior. As explained in Brunswik's Lens Model (Brunswik, 1955), social judgment integrates objective and subjective components. Objective cues in the environment are filtered through a perceiver's subjective lens. Perceivers may attend to some cues and ignore others in making their judgments. Thus, the gender difference in how women and men leaders were perceived could arise not only from how the leaders behaved, but also from how participants interpreted the valence of their behavior. Understanding the cumulative target and perceiver effects provides valuable information about how people perceive the pattern available to them. Nonetheless, it will be important for future work to disentangle perceiver and target components. Such explorations could involve computer identification of emotion, though these can also be biased by gender (Chen & Joo, 2021), or controlling for participants' gender stereotypes of emotion.

Cultural Influence of Nonverbal Bias: Implications and Limitations

Overall, and specifically concerning the consistent effects on White leaders, this work supports a growing body of research suggesting that people can extract complex patterns from subtle cues in their environments (Camp et al., 2021; Lamer et al., 2022; Lamer & Weisbuch, 2019; Miyamoto et al., 2006; Weisbuch et al., 2009; Weisbuch & Ambady, 2009). The cultural snapshots approach (Weisbuch et al., 2017) allowed us to test how people can learn using stimuli with more visual complexity than static expressions or a single interaction. Perceivers had to extract the pattern across clips, settings, and target characters. Natural social environments typically contain a variety of cues, rendering nonverbal behavior just one among many that could be extracted by a participant and inform their beliefs. Nonetheless, participants could extract meaning from patterns presented in this way. The current work contributes to the literature on gender bias in leadership, proposing a mechanism that could be applied to understanding how bias can be socialized across a large population, even when explicit statements claiming support for women being leaders are normative (Parker et al., 2018).

Considering Effect Sizes. The effect sizes in the current work are small, and the experimental manipulation used the most positive and negative clips sampled. In the 361 clips from Study 1, women leaders were perceived as being treated more negatively than men leaders and perceived to have acted more negatively than men leaders. The subset of 16 Pro-Man clips shown to participants in Studies 2 and 3 amplified the pattern people likely encounter when they watch popular television. Conclusions about how nonverbal bias shapes gendered leadership preferences should be considered with this limitation in mind since many factors shape gender bias in leadership (e.g., Heilman, 2012; Hill et al., 2016; Hoyt, 2010; Morgenroth et al., 2020; Williams & Tiedens,

2016). Nonetheless, the current work is consistent with theorizing that nonverbal bias is one possible mechanism for people to learn social beliefs within a culture (Weisbuch & Pauker, 2011). The shows sampled represent a pattern that a broad audience was exposed to. For example, the 2015 season finale of Grey's Anatomy was watched by 8.1 million viewers when it aired (Mitovich, 2016) and, given streaming services, continues to be available for viewers. Furthermore, participants watched only 16 clips in our experiments, totaling about 3 min. If this brief manipulation was sufficient to shape people's immediate leader preferences, exposure to such a pattern of nonverbal bias accumulated across days, weeks, and months could exert a stronger effect, though such a hypothesis has not yet been tested. Exploring the impact of cumulative exposure to nonverbal bias is an important empirical question for future work.

Considering Intersections With Race. In Studies 2 and 3, nonverbal bias consistently impacted preferences for White men vs. White women, but effects were inconsistent when choosing who to work for among Black women and men bosses. In Study 2, participants showed a similar preference for men leaders over women leaders in the Pro-Man Leader (vs. Pro-Woman Leader) condition regardless of the race of the leaders they were selecting from (i.e., Black or White leaders). In Study 3, the exhibited preference for men leaders was present only for White leaders: people in the Pro-Man condition were less likely to select a woman than were those in the Pro-Woman condition *if* the leader choices were White.

Our data cannot explain why the effect on Black leaders was inconsistent. One possible explanation for this inconsistency is that the manipulation contained fewer clips of Black than White leaders (i.e., four clips relative to 12 clips). In our manipulation, 25% of clips featured Black leaders and 75% of clips featured White leaders. The inconsistent results between Studies 2 and 3 could suggest that a small number of clips embedded in a larger set is insufficient to consistently influence behavior, especially when those clips convey a belief that does not have as strong of a cultural consensus (e.g., the beliefs that Black men are good leaders vs. White men are good leaders; Wilson et al., 2017).

Past work employing these thin slices to manipulate beliefs has used 17 (Weisbuch & Ambady, 2009), 22 (Weisbuch et al., 2009), and 24 (Lamer et al., 2022) clips and has targeted a single belief (e.g., attitudes toward Black people or slim women). It remains to be clarified if just a few clips are sufficient to produce these effects among adults or if the effects depend on viewing a *pattern* of nonverbal bias across many clips. This remains an important question for future research to better understand the mechanisms through which nonverbal bias impacts people's beliefs.

Furthermore, this work highlights important future questions at the intersection of gender and race. What messages are being conveyed about leaders varying in gender and race? In Study 2, Black women were selected as leaders a remarkably low proportion of times among participants assigned to the Pro-Man Leader condition, suggesting they may have been punished more so by the pattern in the clips featuring negative hierarchical interactions with women leaders. Yet, this effect disappeared in Study 3, calling the validity of this interpretation into question. Although answering these questions is beyond the scope of these data, future better-powered studies should weigh in on the generalizability of this effect across race, testing nonverbal bias effects through an intersectional lens.

Considering Ingroup Biases and Social Desirability on Leader Preference. Interestingly, although White women leaders were relatively disadvantaged when participants had watched the culturally prevalent Pro-Man Leader pattern relative to the Pro-Woman leader pattern, participants preferred women across both conditions. In other words, more than 50% of participants chose to join a team led by a White woman regardless of condition (55% and 71%). This pattern is consistent with literature quantifying gender ingroup biases. For example, women recruiters tend to exhibit biases favoring job candidates who are also women (Carlsson & Eriksson, 2019), women and men prefer a leader of their own gender more after being reminded of their mortality (Study 1 in Hoyt et al., 2009), and women tend to automatically evaluate women more positively than men (Rudman & Goodwin, 2004). The effects of participant gender were marginal in both experimental studies such that women were non-significantly more likely to choose a woman leader than men, and White women were especially likely to select a White woman. Thus, this overall preference for White women leaders in our sample could reflect an ingroup bias.

Another explanation for the tendency to choose a White but not a Black woman more than 50% of the time is that people may have been more likely to select a woman or man depending on what seemed socially desirable. When selecting between White women and White men, norms would suggest that to appear unbiased, people should select a woman, and indeed, people frequently did. When selecting between Black women and Black men, it is less clear what response from participants would signal being unbiased, as Black women and Black men are both underrepresented in leadership roles and frequently subject to discrimination (Livingston & Pearce, 2009).

Considering Verbal Behavior. The current work isolated nonverbal behavior, using silent clips in all studies. Yet the stimuli that people typically encounter contain both nonverbal and verbal information. Future work should extend these findings to understand how nonverbal and verbal behavior interact. For example, do the scripts reflect the same

pattern of bias as the nonverbal behavior? Are interactions between women leaders and their subordinates more negative than those between men leaders and their subordinates? It will also be valuable to tease apart how verbal and nonverbal cues each impact perceivers. Do they act synergistically? Is one source more influential? Does one kind of cue take precedence? Limited past work has pitted nonverbal and verbal behavior against each other. However, in those studies, nonverbal behavior has emerged above and beyond speech content as the significant predictor of behavior (e.g., Castelli et al., 2008). Testing how verbal and nonverbal cues interact is an important empirical question for future work, especially as researchers seek to understand the causal influence of culturally prevalent patterns.

Practice Implications

This work points to important implications for how media is both created and consumed. When creating media, people wield power to shape stereotypes via subtle patterns of nonverbal behavior embedded in the media they create. Nonverbal behavioral patterns are often hard to detect (Weisbuch & Ambady, 2009; Weisbuch et al., 2009), rendering them difficult to counterargue in the moment. Taken alongside past work documenting numerous other impactful patterns of nonverbal bias in popular media (Lamer et al., 2018; Lamer et al., 2022; Weisbuch & Ambady, 2009; Weisbuch et al., 2009), it is clear that people are exposed to many patterns when they consume media. Thus, this work suggests a critical eye toward media consumption *and* creation.

Conclusion

In this work, we found evidence of a prevalent pattern of nonverbal bias in televised media that favors men leaders over women leaders. In two studies – one exploratory and one confirmatory—exposure to this bias led people to be more likely to show a preference for working with White men leaders than White women leaders—regardless of their own gender. These findings suggest that bias against women leaders exists in cultural patterns of nonverbal behavior and observing such patterns can cause people to discriminate against women leaders.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Sarah Ariel Lamer https://orcid.org/0000-0003-3597-0958
Haley B. Beck https://orcid.org/0000-0003-1820-8398
Leanne ten Brinke https://orcid.org/0000-0001-7080-9129
Gillian Michelle Preston https://orcid.org/0009-0004-6784-5603

Supplemental Material

Supplemental material for this article is available online.

Notes

1. We used Correll and colleagues' (2022) *t*back method to calculate the effect size, eta squared.
2. Note that in both experimental studies, participants could have indicated more than one of these options, thus percentages total more than 100.
3. Because interaction partner and target positivity were separate measures, we explain how we weighed each of these in the clip selection process. Specifically, we sorted the hierarchical clips in ascending order by interaction partner positivity within each show. We then selected the most positive and most negative clip of each leader in a hierarchical interaction based on the interaction partner positivity, confirming that the target behavior was similar (and at the least not more than 1.75 points away on the scale in a way that would be inconsistent with the condition). For example, the hierarchical clip in which the leader from Superstore was treated most positively (4.87) had a target positivity rating of 2.77 (2.10 more negative than the interaction partners). Therefore, we excluded this clip and used the next most positive clip (interaction partner positivity = 4.59, target positivity = 5.68). A table of the final clips used in the manipulation and their nonverbal positivity ratings are available in the Supplementary Online Materials (Table S2). Because participants would only get to see a small subset of clips, we wanted to ensure that each clip clearly contained the emotion and context intended. Thus, we screened the selected clips to ensure that they did not contain cues that would render it difficult for participants to understand the hierarchical nature of characters in the clips (e.g., the leader interacted with subordinates after work while drinking), they did not contain lighting issues that would make emotion perception tricky (e.g., very dark rooms), and that they did not contain highly ambivalent emotion from interaction partners (e.g., one interaction partner being very positive and one interaction partner being very negative). Finally, within each show, we ensured that the average difference between the emotions of the selected clips was in the direction expected (e.g., the leader and subordinates in clips featuring the man leader from Superstore were more positive in the Pro-Man than Pro-Woman condition).
4. When told at the end of the study after completing all measures that there was a pattern in the clips and asked to identify it, between 9% (Study 3) and 15% (Study 2) of participants mentioned gender in some way (e.g., "The gender of the leaders changed every two videos," "women leader," "woman

leading convos") with the majority of participants guessing something about the order of clips featuring each gender (e.g., alternating) but no one accurately guessing the pattern involving emotion in the clips they saw.

5. As part of a project on measurement construction, we also included a variety of gendered traits ($N = 48$) to test the factor structure of gender stereotypes and a leadership styles questionnaire (ten Brinke & Keltner, 2020). We did not observe any effects of condition on these other measures.
6. We also included an exploratory measure of participant desire to hold a leadership position (Davies et al., 2005) and gender identity importance (Luhtanen & Crocker, 1992). There were no main or interactive effects of condition on one's own leadership interests. See a description of this outcome and its results in the Supplementary Materials.

References

Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018). Journal article reporting standards for quantitative research in psychology: The APA Publications and Communications Board task force report. *American Psychologist*, *73*(1), 3–25. https://doi.org/10.1037/amp0000191

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bogaard, G., Meijer, E. H., Vrij, A., & Merckelbach, H. (2016). Strong, but wrong: Lay people's and police officers' beliefs about verbal and nonverbal cues to deception. *PLoS ONE*, *11*(6), 1–19. https://doi.org/10.1371/journal.pone.0156615

Bongiorno, R., Bain, P. G., & David, B. (2014). If you're going to be a leader, at least act like it! Prejudice towards women who are tentative in leader roles. *British Journal of Social Psychology*, *53*(2), 217–234. https://doi.org/10.1111/bjso.12032

Bosak, J., & Sczesny, S. (2011). Gender bias in leader selection?: Evidence from a hiring simulation study. *Sex Roles*, *65*(3–4), 234–242. https://doi.org/10.1007/s11199-011-0012-7

Brey, E., & Pauker, K. (2019). Teachers' nonverbal behaviors influence children's stereotypic beliefs. *Journal of Experimental Child Psychology*, *188*(104671), 1–13. https://doi.org/10.1016/j.jecp.2019.104671

Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review*, *62*(3), 193–217. https://doi.org/10.1037/h0047470

Camp, N. P., Voigt, R., Jurafsky, D., & Eberhardt, J. L. (2021). The thin blue waveform: Racial disparities in officer prosody undermine institutional trust in the police. *Journal of Personality & Social Psychology*, *121*(6), 1157–1171. https://doi.org/10.1037/pspa0000270

Carlsson, M., & Eriksson, S. (2019). In-group gender bias in hiring: Real-world evidence. *Economics Letters*, *185*(108686), 1–3. https://doi.org/10.1016/j.econlet.2019.108686

Castelli, L., De Dea, C., & Nesdale, D. (2008). Learning social attitudes: Children's sensitivity to the nonverbal behaviors of adult models during interracial interactions. *Personality and Social Psychology Bulletin*, *34*(11), 1504–1513. https://doi.org/10.1177/0146167208322769

Catalyst. (2019). Historical list of women CEOs of the Fortune lists: 1972-2019.

Center for American Women and Politics. (2024). Women in the U.S. Senate 2024. https://cawp.rutgers.edu/facts/levels-office/congress/women-us-senate-2024.

Chambers, C. D., Feredoes, E., Muthukumaraswamy, S. D., & Etchells, P. J. (2014). Instead of "playing the game" it is time to change the rules: Registered Reports at AIMS Neuroscience and beyond. *AIMS Neuroscience*, *1*(1), 4–17. https://doi.org/10.3934/Neuroscience.2014.1.4

Chen, Y., & Joo, J. (2021). Understanding and mitigating annotation bias in facial expression recognition. IEEE/CVF International Conference on Computer Vision (ICCV), 14960–14971. https://doi.org/10.1109/ICCV48922.2021.01471

Correll, J., Mellinger, C., & Pedersen, E. J. (2022). Flexible approaches for estimating partial eta squared in mixed-effects models with crossed random factors. *Behavior Research Methods*, *54*, 1626–1642. https://doi.org/10.3758/s13428-021-01687-2

Curtis, N., & Cardo, V. (2018). Superheroes and third-wave feminism. *Feminist Media Studies*, *18*(3), 381–396. https://doi.org/10.1080/14680777.2017.1351387

Davies, P. G., Steele, C. M., & Spencer, S. J. (2005). Clearing the air: Identity safety moderates the effects of stereotype threat on women's leadership aspirations. *Journal of Personality and Social Psychology*, *88*(2), 276–287. https://doi.org/10.1037/0022-3514.88.2.276

Delmas, E., Besson, M., Brice, M. H., Burkle, L. A., Dalla Riva, G. V., Fortin, M. J., Gravel, D., Guimarães, P. R., Hembry, D. H., Newman, E. A., Olesen, J. M., Pires, M. M., Yeakel, J. D., & Poisot, T. (2019). Analysing ecological networks of species interactions. *Biological Reviews*, *94*(1), 16–36. https://doi.org/10.1111/brv.12433

DePaulo, B. M. (1992). Nonverbal behavior and self-presentation. *Psychological Bulletin*, *111*(2), 203–243. https://doi.org/10.1037/0033-2909.111.2.203

Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review*, *109*(3), 573–598. https://doi.org/10.1037/0033-295X.109.3.573

Eitan, O., Viganola, D., Inbar, Y., Dreber, A., Johannesson, M., Pfeiffer, T., Thau, S., & Uhlmann, E. L. (2018). Is research in social psychology politically biased?: Systematic empirical tests and a forecasting survey to address the controversy. *Journal of Experimental Social Psychology*, *79*(November), 188–199. https://doi.org/10.1016/j.jesp.2018.06.004

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social,

behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Fleischer, M., & Fleischer, D. (1941). *Superman [Motion picture]*. Fleischer Studios.

Gelman, A., & Carlin, J. (2014). Beyond power calculations: Assessing type s (sign) and type m (magnitude) errors. *Perspectives on Psychological Science*, *9*(6), 641–651. https://doi.org/10.1177/1745691614551642

Giner-Sorolla, R., Montoya, A. K., Reifman, A., Carpenter, T., Lewis, N. A., Aberson, C. L., Bostyn, D. H., Conrique, B. G., Ng, B. W., Schoemann, A. M., & Soderberg, C. (2024). Power to detect what? Considerations for planning and evaluating sample size. *Personality and Social Psychology Review*, *28*(3), 276–301. https://doi.org/10.1177/10888683241228328

Green, P., & Macleod, C. J. (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, *7*(4), 493–498. https://doi.org/10.1111/2041-210X.12504

Heilman, M. E. (2012). Gender stereotypes and workplace bias. *Research in Organizational Behavior*, *32*, 113–135. https://doi.org/10.1016/j.riob.2012.11.003

Hill, C., Miller, K., Benson, K., & Handley, G. (2016). Barriers and bias: The status of women in leadership. American Association of University Women.

Hinchliffe, E. (2024, June 4). The share of Fortune 500 companies run by women CEOs stays flat at 10.4% as pace of change stalls. Fortune. https://fortune.com/2024/06/04/fortune-500-companies-women-ceos-2024/.

Hoewe, J., & Sherrill, L. A. (2019). The influence of female lead characters in political TV shows: Links to political engagement. *Journal of Broadcasting and Electronic Media*, *63*(1), 59–76. https://doi.org/10.1080/08838151.2019.1570782

Hoyt, C. L. (2010). Women, men, and leadership: Exploring the gender gap at the top. *Social and Personality Psychology Compass*, *4*(7), 484–498. https://doi.org/10.1111/j.1751-9004.2010.00274.x

Hoyt, C. L., Simon, S., & Reid, L. (2009). Choosing the best (wo)man for the job: The effects of mortality salience, sex, and gender stereotypes on leader evaluations. *Leadership Quarterly*, *20*(2), 233–246. https://doi.org/10.1016/j.leaqua.2009.01.016

Jarvis, B. (2012). MediaLab (2012.4.0.166). Empirisoft Corporation.

Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, *103*(1), 54–69. https://doi.org/10.1037/a0028347

Koenig, A. M., Eagly, A. H., Mitchell, A. A., & Ristikari, T. (2011). Are leader stereotypes masculine? A meta-analysis of three research paradigms. *Psychological Bulletin*, *137*(4), 616–642. https://doi.org/10.1037/a0023557

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Lamer, S. A., Dvorak, P., Biddle, A. M., Pauker, K., & Weisbuch, M. (2022). The transmission of gender stereotypes through televised patterns of nonverbal bias. *Journal of Personality and Social Psychology*, *123*(6), 1315–1335. https://doi.org/10.1037/pspi0000390

Lamer, S. A., Sweeny, T. D., Dyer, M. L., & Weisbuch, M. (2018). Rapid visual perception of interracial crowds: Racial category learning from emotional segregation. *Journal of Experimental Psychology: General*, *147*(5), 683–701. https://doi.org/10.1037/xge0000443

Lamer, S. A., & Weisbuch, M. (2019). Men over women: The social transmission of gender stereotypes through spatial elevation. *Journal of Experimental Social Psychology*, *84*(103828), 1–19. https://doi.org/10.1016/j.jesp.2019.103828

Livingston, R. W., & Pearce, N. A. (2009). The teddy-bear effect: Does having a baby face benefit black chief executive officers? *Psychological Science*, *20*(10), 1229–1236. https://doi.org/10.1111/j.1467-9280.2009.02431.x

Livingston, R. W., Rosette, A. S., & Washington, E. F. (2012). Can an agentic Black woman get ahead? The impact of race and interpersonal dominance on perceptions of female leaders. *Psychological Science*, *23*(4), 354–358. https://doi.org/10.1177/0956797611428079

Luhtanen, R., & Crocker, J. (1992). A collective self-esteem scale: Self-evaluation of one's social identity. *Personality and Social Psychology Bulletin*, *18*(3), 302–318. https://doi.org/10.1177/0146167292183006

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods*, *47*(4), 1122–1135. https://doi.org/10.3758/s13428-014-0532-5

Maher, J. M., Markey, J. C., & Ebert-May, D. (2013). The other half of the story: Effect size analysis in quantitative research. *CBE Life Sciences Education*, *12*(3), 345–351. https://doi.org/10.1187/cbe.13-04-0082

Mastro, D. E., & Greenberg, B. S. (2000). The portrayal of racial minorities on prime time television. *Journal of Broadcasting and Electronic Media*, *44*(4), 690–703. https://doi.org/10.1207/s15506878jobem4404_10

McGregor, J. (2014, November 5). More than 100 women in Congress for the first time, but not much growth. The Washington Post. https://www.washingtonpost.com/news/on-leadership/wp/2014/11/05/more-than-100-women-in-congress-for-the-first-time-but-not-much-growth/.

Mitovich, M. W. (2016, May 20). Ratings: Grey's Anatomy up from previous finale, hits 14-week highs. TV Line. https://tvline.com/ratings/greys-anatomy-season-12-finale-ratings-716255/.

Miyamoto, Y., Nisbett, R. E., & Masuda, T. (2006). Culture and the physical environment holistic versus analytic perceptual affordances. *Psychological Science*, *17*(2), 113–119. https://doi.org/10.1111/j.1467-9280.2006.01673.x

Monk-Turner, E., Heiserman, M., & Johnson, C. (2010). The portrayal of racial minorities on prime time television : A replication of the mastro and greenberg study a decade later. *Studies in Popular Culture*, *32*(2), 101–114. https://doi.org/10.1207/s15506878jobem4404_10

Morgenroth, T., Kirby, T. A., Ryan, M. K., & Sudkämper, A. (2020). The who, when, and why of the glass cliff

phenomenon: A meta-analysis of appointments to precarious leadership positions. *Psychological Bulletin*, *146*(9), 797–829. https://doi.org/10.1037/bul0000234

Oppenheimer, B., Goodman, M., Adams-Price, C., Codling, J., & Coker, J. D. (2003). Audience perceptions of strong female characters on television. *Communication Research Reports*, *20*(2), 161–172. https://doi.org/10.1080/08824090309388812

Parker, K., Horowitz, J., & Igielnik, R. (2018, September 20). Women and Leadership 2018. Pew Research Center. https://www.pewresearch.org/social-trends/2018/09/20/women-and-leadership-2018/.

Petsko, C. D., & Bodenhausen, G. V. (2019). Racial stereotyping of gay men: Can a minority sexual orientation erase race? *Journal of Experimental Social Psychology*, *83*(2019), 37–54. https://doi.org/10.1016/j.jesp.2019.03.002

Phelan, J. E., Moss-Racusin, C. A., & Rudman, L. A. (2008). Competent yet out in the cold: Shifting criteria for hiring reflect backlash toward agentic women. *Psychology of Women Quarterly*, *32*(4), 406–413. https://doi.org/10.1111/j.1471-6402.2008.00454.x

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/.

Rosette, A. S., Koval, C. Z., Ma, A., & Livingston, R. (2016). Race matters for women leaders: Intersectional effects on agentic deficiencies and penalties. *Leadership Quarterly*, *27*(3), 429–445. https://doi.org/10.1016/j.leaqua.2016.01.008

Rudman, L. A. (1998). Self-promotion as a risk factor for women: The costs and benefits of counterstereotypical impression management. *Journal of Personality and Social Psychology*, *74*(3), 629–645. https://doi.org/10.1037/0022-3514.74.3.629

Rudman, L. A., & Glick, P. (2001). Prescriptive gender stereotypes and backlash toward agentic women. *Journal of Social Issues*, *57*(4), 743–762. https://doi.org/10.1111/0022-4537.00239

Rudman, L. A., & Goodwin, S. A. (2004). Gender differences in automatic in-group bias: Why do women like women more than men like men? *Journal of Personality and Social Psychology*, *87*(4), 494–509. https://doi.org/10.1037/0022-3514.87.4.494

Saha, S., & Weeks, A. C. (2022). Ambitious women: Gender and voter perceptions of candidate ambition. *Political Behavior*, *44*(2), 779–805. https://doi.org/10.1007/s11109-020-09636-z

Sánchez-Meca, J., Marín-Martínez, F., & Chacón-Moscoso, S. (2003). Effect-size indices for dichotomized outcomes in meta-analysis. *Psychological Methods*, *8*(4), 448–467. https://doi.org/10.1037/1082-989X.8.4.448

Schwarz, S., & Coppock, A. (2022). What have we learned about gender from candidate choice experiments? A meta-analysis of sixty-seven factorial survey experiments. *Journal of Politics*, *84*(2), 655–668. https://doi.org/10.1086/716290

Sherman, A., Sweeny, T. D., Grabowecky, M., & Suzuki, S. (2012). Laughter exaggerates happy and sad faces depending on visual context. *Psychonomic Bulletin and Review*, *19*(2), 163–169. https://doi.org/10.3758/s13423-011-0198-2

Signorielli, N., & Bacue, A. (1999). Recognition and respect: A content analysis of prime-time television characters across three decades. *Sex Roles*, *40*(7–8), 527–544. https://doi.org/10.1023/A:1018883912900

Sink, A., & Mastro, D. (2017). Depictions of gender on primetime television: A quantitative content analysis. *Mass Communication and Society*, *20*(3), 3–22. https://doi.org/10.1080/15205436.2016.1212243

Skinner, A. L., Meltzoff, A. N., & Olson, K. R. (2017). "Catching" social bias: Exposure to biased nonverbal signals creates social biases in preschool children. *Psychological Science*, *28*(2), 216–224. https://doi.org/10.1177/0956797616678930

Smith, S. L., & Granados, A. D. (2009). Content patterns and effects surrounding sex-role stereotyping on television and film. In J. Bryant & M. B. Oliver (Eds.), *Media effects: Advances in theory and research* (3rd ed, pp. 342–361). Routledge.

Smith, J. L., Paul, D., & Paul, R. (2007). No place for a woman: Evidence for gender bias in evaluations of presidential candidates. *Basic and Applied Social Psychology*, *29*(3), 225–233. https://doi.org/10.1080/01973530701503069

Smith, S. L., Pieper, K. M., Granados, A. D., & Choueiti, M. (2010). Assessing gender-related portrayals in top-grossing g-rated films. *Sex Roles*, *62*(11), 774–786. https://doi.org/10.1007/s11199-009-9736-z

Stoll, J. (2021). *Weekly TV Reach in the United States as of the 1st Quarter of 2019, by Age Group*. Statista.

Stoll, J. (2022). *Average Daily Time Spent Watching TV in the United States from 2019 to 2023*. Statista.

Teele, D. L., Kalla, J., & Rosenbluth, F. (2018). The ties that double bind: Social roles and women's underrepresentation in politics. *American Political Science Review*, *112*(3), 525–541. https://doi.org/10.1017/S0003055418000217

ten Brinke, L., & Keltner, D. (2020). Theories of power: Perceived strategies for gaining and maintaining power. *Journal of Personality and Social Psychology*, *122*(1), 53–72. https://doi.org/10.1037/pspi0000345

Vial, A. C., & Napier, J. L. (2018). Unnecessary frills: Communality as a nice (but expendable) trait in leaders. *Frontiers in Psychology*, *9*, 1–15. https://doi.org/10.3389/fpsyg.2018.01866

Weisbuch, M., & Ambady, N. (2009). Unspoken cultural influence: Exposure to and influence of nonverbal bias. *Journal of Personality & Social Psychology*, *96*(6), 1104–1119. https://doi.org/10.1037/a0015642

Weisbuch, M., Lamer, S. A., Treinen, E., & Pauker, K. (2017). Cultural snapshots: Theory and method. *Social and Personality Psychology Compass*, *11*(9), 1–21. https://doi.org/10.1111/spc3.12334

Weisbuch, M., & Pauker, K. (2011). The nonverbal transmission of intergroup bias: A model of bias contagion with implications for social policy. *Social Issues and Policy Review*, *5*(1), 257–291. https://doi.org/10.1038/jid.2014.371

Weisbuch, M., Pauker, K., & Ambady, N. (2009). The subtle transmission of race bias via televised nonverbal behavior. *Science*, *326*(5960), 1711–1714. https://doi.org/10.1126/science.1178358

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag. ISBN 978-3-319-24277-4, https://ggplot2.tidyverse.org.

Williams, M. J., & Tiedens, L. Z. (2016). The subtle suspension of backlash: A meta-analysis of penalties for women's implicit and explicit dominance behavior. *Psychological Bulletin*, *142*(2), 165–197. https://doi.org/10.1037/bul0000039

Wilson, J. P., Remedios, J. D., & Rule, N. O. (2017). Interactive effects of obvious and ambiguous social categories on perceptions of leadership: When double-minority status may be beneficial. *Personality and Social Psychology Bulletin*, *43*(6), 888–900. https://doi.org/10.1177/0146167217702373