



Hiring women into senior leadership positions is associated with a reduction in gender stereotypes in organizational language

M. Asher Lawson^{a,1}, Ashley E. Martin^b, Imrul Huda^c, and Sandra C. Matz^d

^aFuqua School of Business, Duke University, Durham, NC 27708; ^bGraduate School of Business, Stanford University, Stanford, CA 94305; ^cTMW Center for Early Learning and Public Health, University of Chicago, Chicago, IL 60637; and ^dColumbia Business School, Columbia University, New York, NY 10027

Edited by Jonah Berger, The Wharton School, University of Pennsylvania, Philadelphia, PA; received December 30, 2020; accepted December 22, 2021 by Editorial Board Member Jennifer A. Richeson

Women continue to be underrepresented in leadership positions. This underrepresentation is at least partly driven by gender stereotypes that associate men, but not women, with achievement-oriented, agentic traits (e.g., assertive and decisive). These stereotypes are expressed and perpetuated in language, with women being described in less agentic terms than men. The present research suggests that appointing women to the top tiers of management can mitigate these deep-rooted stereotypes that are expressed in language. We use natural language processing techniques to analyze over 43,000 documents containing 1.23 billion words, finding that hiring female chief executive officers and board members is associated with changes in organizations' use of language, such that the semantic meaning of being a woman becomes more similar to the semantic meaning of agency. In other words, hiring women into leadership positions helps to associate women with characteristics that are critical for leadership success. Importantly, our findings suggest that changing organizational language through increasing female representation might provide a path for women to break out of the double bind: when female leaders are appointed into positions of power, women are more strongly associated with the positive aspects of agency (e.g., independent and confident) in language but not at the cost of a reduced association with communality (e.g., kind and caring). Taken together, our findings suggest that female representation is not merely an end, but also a means to systemically change insidious gender stereotypes and overcome the trade-off between women being perceived as either competent or likeable.

stereotypes | language | gender inequality

At all levels of management, women continue to be underrepresented (1). Nowhere is this more pronounced than at the top levels of leadership, including chief executive officer (CEO) positions and board memberships. This underrepresentation is thought to be due, at least in part, to gender stereotypes, where men are associated with achievement-oriented, agentic traits (e.g., assertive and decisive) and women are not (2, 3). Agentic traits are those valued and required in positions of leadership, creating the expectation that men will be more likely to possess the skills required for occupational success (3, 4). These gender stereotypes lead to a double bind for women. Women are either considered less agentic and therefore less competent than men, or they are seen as equally agentic but less likeable because they violate gender stereotypes (2–6). The consequences of such gender stereotypes have been well documented, resulting in the devaluation of women's performance, denial of credit for their success, and exclusion in the workplace (5, 6).

Prior research has examined numerous ways to minimize or dissolve gender stereotypes that associate men—but not women—with agency (7–11). Such interventions range from training programs (7) and diversity education (9, 11) to intentional habit formation (8, 10). However, the efficacy of formal training programs and education has been mixed (7, 8, 10).

A promising but underexplored avenue for overcoming harmful gender stereotypes is to change the way that organizations express these stereotypes in the language they use to describe women. An organization's use of language can both reflect and perpetuate gender stereotypes, such as the stereotype that women are less agentic than men (2, 12–14). Research has shown that this gendering of language can be a major roadblock on the path to greater gender parity in organizations (2). For example, the gendered language in job advertisements negatively influences women's likelihood of applying for high-status positions (15), and the language used in promotion letters affects women's ability to climb the career ladder at the same speed as their male counterparts (13, 16, 17).

How can organizations change their language such that women are described in more agentic terms? How can this shift be implemented in a way that allows women to break out of the double bind to be seen as more agentic but no less warm and likeable? In practice, changing language on an individual level is difficult because it requires cognitive effort and resources (18, 19), which might explain why previous efforts have fallen short (7). Even when women are described in more agentic terms, they are often characterized by the more negative aspects of agency [e.g., being dominant or aggressive (6, 20)] and face social and economic backlash for violating traditional

Significance

Gender inequality has been deemed the “greatest human rights challenge of our time” by the United Nations, and scholars across numerous disciplines agree that gender stereotypes represent a primary way by which this inequality is maintained. Yet changing stereotypes in a systemic, enduring way is extremely difficult. This is at least in part because stereotypes are transmitted and perpetuated through the language societies and organizations use to describe women, especially those in leadership roles. Here, we show that hiring women into leadership positions is associated with organizations characterizing women in more leadership-congruent, agentic ways. This shift mitigates a critical barrier to women's progression in organizations and society: the incongruence of what it means to be a woman and a leader.

Author contributions: M.A.L., A.E.M., and S.C.M. designed research; M.A.L. and S.C.M. performed research; M.A.L. and I.H. analyzed data; and M.A.L., A.E.M., and S.C.M. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission. J.B. is a guest editor invited by the Editorial Board.

This open access article is distributed under [Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 \(CC BY-NC-ND\)](https://creativecommons.org/licenses/by-nc-nd/4.0/).

¹To whom correspondence may be addressed. Email: asher.lawson@duke.edu.

This article contains supporting information online at <http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2026443119/-/DCSupplemental>.

Published February 22, 2022.

gender norms that prescribe women to be warm and communal (5, 6). This paradox presents a conundrum as the very behaviors that women must engage in to be successful leaders are described in negative, undesirable terms when performed by them (6). Thus, interventions aimed at changing the gendering of language are most likely to succeed if they systemically impact the entire organization, rather than placing the burden of stereotype change on individuals, and if they provide positive and successful examples of agentic, female leadership.

In this paper we show that hiring women into leadership positions—as CEOs or members of executive boards—can provide pathways to systemically and positively changing gender stereotypes encoded in language. We suggest two processes by which increased female representation could induce systemic and positive language change: (1) unconscious, incidental shifts in attitudes toward the appointed women and/or (2) conscious, strategic intention to signal the agentic qualities of organizations' female leaders.

By occupying traditionally male-dominated positions, female exemplars might change gender stereotypes by leading organizational members to associate those women with agentic, leadership behaviors that were previously performed by men (21). This shift in the association of women with agentic behaviors might gradually turn into associations between these women and broader agentic traits that are not tied to specific actions. For example, if employees repeatedly see their female CEO making decisions, leading conversations, and taking initiative, they might come to associate her with decisiveness, confidence, and ambition. While these changes would start with the appointed women, over time the association between women and agentic traits may permeate the organization more widely, with women in general becoming more closely associated with agency. In line with this proposition, prior work has shown that placing female role models in agentic domains (e.g., science and management) can alter gender stereotypes about women's competency and fit in those environments (21). Such shifts in stereotypes occur for explicit stereotyping but also on an implicit level that lies below conscious awareness and shapes our perceptions by mere association (21, 22). Consequently, we expect attitude changes—whether toward the appointed women or women in general—to be incidentally reflected in the language used to describe women.

Yet, if language change was purely incidental, we might also expect it to reflect the aforementioned double bind for women (5, 6). While reading or hearing about women performing agentic tasks might lead people to think of them as more agentic in the long run, it might also lead them to think of those women—and women in general—as less warm and likeable. A woman making tough decisions about downsizing or aggressively expanding the company's market share might lead to perceptions of her being decisive and confident but also cold and callous. Thus, purely incidental changes in the gendering of organizational language might carry unintended consequences that could undermine the benefits of women being perceived and described in more leadership-congruent terms.

However, we argue that it is unlikely that organizational language will reinforce the double bind that women typically face. Organizations that appoint women to the top tiers of leadership are incentivized to strategically convey the positive qualities of their female leaders. That is, they should be motivated to make them appear both agentic and likeable, thereby increasing the likelihood of women being able to break out of the double bind in how they are characterized in organizational language. This argument is in line with signaling theory, which suggests that organizations use language to signal their beliefs, values, and legitimacy to attract employees, consumers, and investors (23, 24). For example, organizations might signal commitment to social responsibility and environmental practices to increase

employee attraction and distinguish themselves from competitors (25, 26). Given that organizations that promote women to leadership positions should be motivated to signal the competence and likeability of their female leaders through the language they use to describe them, a signaling theory account would expect an increase to women's agency, while avoiding the negative effect on warmth and likeability that is typically associated with it.

One important way that organizations communicate shifts in the gendering of language—both internally and externally—is through organizational documents, such as shareholder reports and investor documents [e.g., ISO 9000 certifications (23)]. Using a variety of public organizational documents—including 43,396 files and 1.23 billion words—we show that hiring women as CEOs or board members is associated with changes in organizations' language such that the semantic meaning of female words (e.g., she, her, and woman) becomes more similar to the semantic meaning of agentic words (e.g., competent, independent, and assertive). Importantly, the change in meaning is not linked with a simultaneous drop in the association between women and communality. Providing additional evidence that female representation might allow women to break out of the double bind, we show that the effect of hiring a female CEO on the semantic association of women with agency is predominantly driven by positive agentic terms (e.g., confident and independent) rather than negative ones (e.g., dominant and aggressive). Overall, we show that the effect of appointing women to leadership positions goes beyond merely increasing the immediate representation of women in organizations. Introducing female leadership exemplars allows organizations to battle insidious gender stereotypes by changing the associations, meaning, and interpretations of women's agentic behavior.

Study 1: Hiring a Female CEO Is Associated with Changes in the Gendering of Language

Study 1 aimed to test whether hiring a female CEO is related to an increase in the association of women with agency between the periods before and after hire (Fig. 1). To capture organizations' outward-facing language, we extracted 1) DEF 14A filings (definitive proxy statements), 2) 10-K filings (annual reports), and 3) transcripts of investor calls for all S&P (Standard and Poor's) 500 organizations in a 10 y window between 2009 and 2018. We identified the subset of these organizations that had 1) a female CEO in the sample period (41 organizations), 2) text data available for at least 3 y pre-hiring and post-hiring, and 3) a male CEO predecessor prior to the female hire. This selection procedure resulted in a total of 11 target organizations. For each of these 11 target organizations, we identified two propensity matched S&P 500 organizations that did not hire a female CEO, allowing us to account for changes in gendered language that were not related to the hiring of a woman as CEO. We also performed this same process for female CEOs who were replaced by male CEOs in the eligible period. This provided two additional target organizations and four propensity matches. In total, we had a main sample of 33 organizations and an additional sample of 6 organizations that we analyzed separately in *SI Appendix*.

We measure the change in the gendering of organizations' language before and after hiring a female CEO using an unsupervised natural language processing algorithm [word2vec (27–29)] that estimates the semantic meaning of references to gender and references to agency. The intuition behind word2vec is reflected in the quote by the linguist John Rupert Firth, “You shall know a word by the company it keeps” (30). Seeing a word across many different contexts (sets of surrounding words) enables the algorithm to learn the word's meaning from

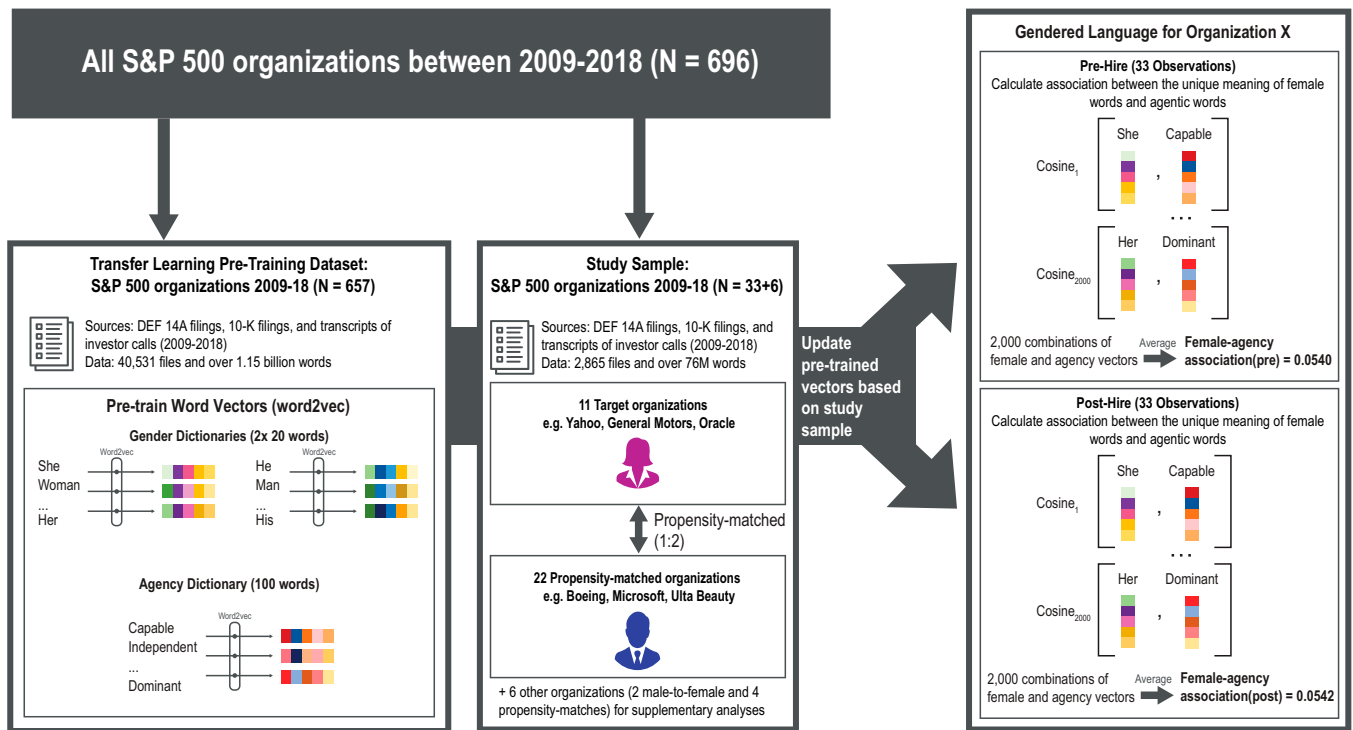


Fig. 1. Procedure for extraction of female-agency and male-agency associations in organizations' public documents in the prehire and posthire periods.

the contexts in which it is used. word2vec can be thought of as analogous to running factor analyses that uncover latent dimensions on which different words vary systematically. The number of these dimensions is specified in the model, but there is no a priori specification of how particular words will relate to or "load on" these dimensions. The output of the word2vec algorithm is a dictionary containing a vector for every word in the text, with a value for every dimension specified.

Due to the complexity of the space, no individual dimension of a word vector is interpretable. However, studying differences between vectors can yield remarkable insights regarding the semantic meaning of words and contribute to the understanding of gender stereotypes (31–33). In fact, the study of analogical reasoning using word vectors has highlighted the inherent gender biases in language, showing, for instance, that historically "man is to computer programmer as woman is to homemaker" (34, p. 1). In the present research, we study the distances between gendered words and agency words in vector space. This allows us to capture the strength of association between what it means to be a woman and what it means to be agentic in organizations' language before and after the hire of a female CEO.

We estimated the female-agency association and male-agency association in the following four steps (see *Methods* and *SI Appendix* for further details): First, we developed and validated an agency dictionary of 100 words (e.g., capable, independent, and dominant) and identified 20 female-related words that could be matched with a direct male counterpart (e.g., she/he, woman/man, and her/his). Second, we used the official text of all 657 organizations that had entered the S&P 500 between 2009 and 2018 but were not in our set of 33 focal organizations (or the 6 other organizations), to pretrain a set of vectors capturing the semantic meaning for each of the words in the two dictionaries. This model was trained based on 12,974 SEC (US Securities and Exchange Commission) documents and 27,557 transcripts with a total of 1.15 billion words. Third, we updated the vectors based on the data of each of the 33 focal organizations separately for the periods before and after hiring the CEO (with an average number of

37 new documents containing 980,000 words per organization period). Finally, we calculated the average cosine similarity between the word vectors representing references to being female or male and the word vectors representing the meaning of agency, for each organization and prehiring/posthiring time period.

We first explored our hypothesis—that the hiring of a female CEO would be associated with a strengthening of the semantic relationship between female and agentic vectors in the target organizations relative to the propensity-matched organizations—in a descriptive way. Specifically, we visualized the grand mean-centered cosine similarity between our female and agency vectors, at an organization level, in the periods before and after the female CEOs were hired (Fig. 2). For example, comparing General Motors—which hired Mary T. Barra in 2014—and Ford Motor Company shows that while Ford saw a drop in the female-agency association after 2014, General Motors saw an increase. Providing initial support for our hypothesis, Fig. 2 shows that 8 of the 11 target organizations (72.7%) saw an increase in the association between the semantic meaning of female and agency words, while only 6 of the 22 control organizations did in the same time period (27.3%).

To test our hypothesis more formally, we used linear regression analysis, regressing either the female-agency association or the male-agency association in organizations' language onto the binary coded time period (prehiring = 0, posthiring = 1) and the organization category (0 = propensity match, 1 = target), as well as their two-way interaction. Our main variable of interest, the interaction term (β_3), indicates the extent to which target organizations experienced a stronger increase in the association between gender words and agency words relative to their propensity-matched peers.

$$FA = \beta_1(\text{period}) + \beta_2(\text{target}) + \beta_3(\text{period} \times \text{target}), \quad [1]$$

$$MA = \beta_1(\text{period}) + \beta_2(\text{target}) + \beta_3(\text{period} \times \text{target}). \quad [2]$$

To increase the robustness of our findings, we used 27 different specifications (3 gender dictionaries, 3 agency dictionaries, and

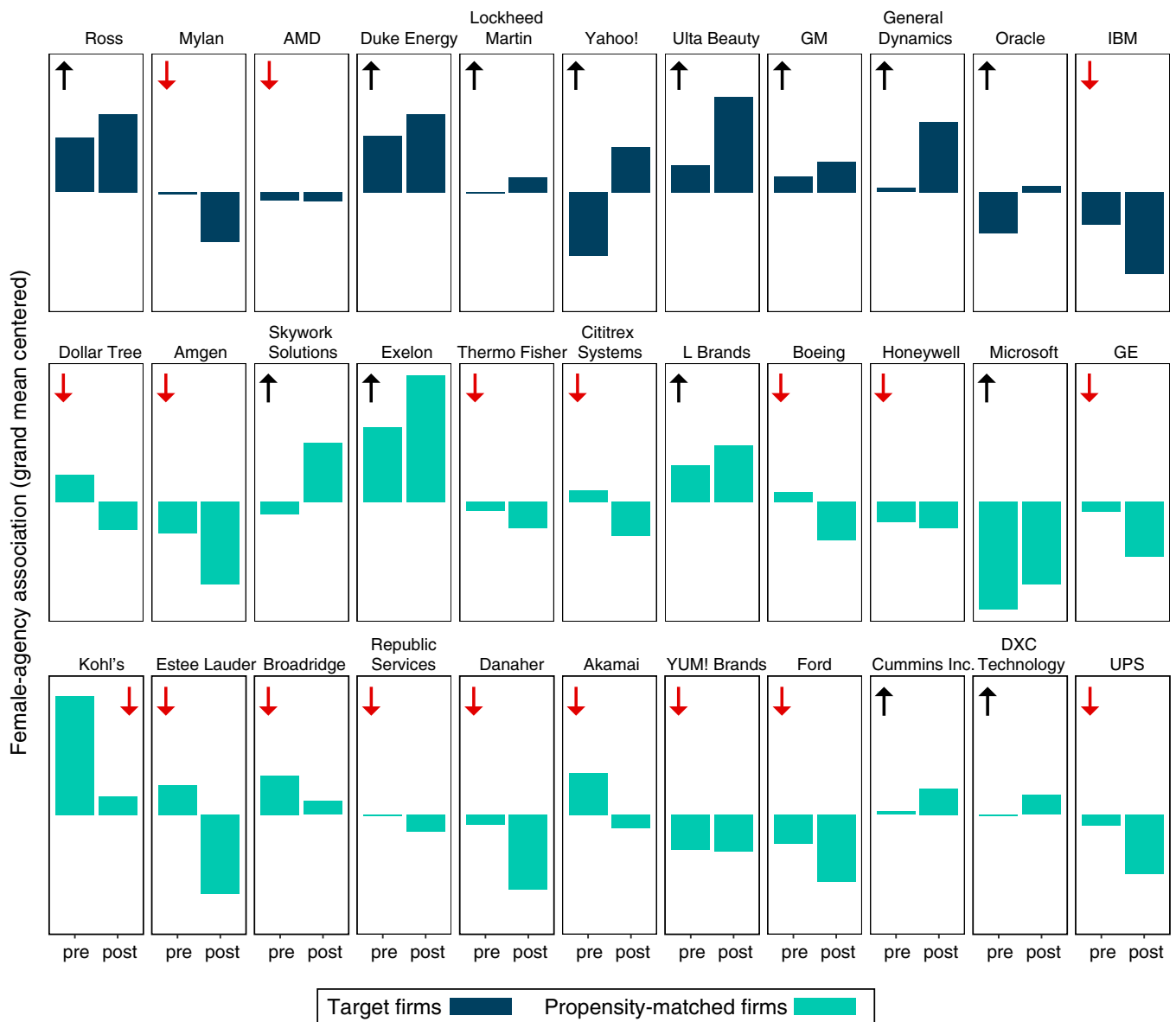


Fig. 2. Average cosine similarity between female vectors and agency vectors for each organization in the periods pre- and post-CEO hire, grand mean centered.

3 dimensions of word vectors) and drew 1,000 bootstrapped samples for each specification from our 11 target and 22 propensity matched organizations, resulting in a total of 27,000 iterations. In Fig. 3, we present the standardized regression coefficients (betas) of the interaction term from the 27,000 model estimates, both for the female-agency and the male-agency effects (ordered by effect size). Providing support for our hypothesis, the findings show that the pre-hiring–post-hiring change in female-agency association was significantly larger for organizations that hired a female CEO [average $\beta_3 = 0.226$, $SD(\beta_3) = 0.076$]. With the vast majority of standardized regression coefficients being positive (99.996%), the distribution of coefficients was significantly larger than zero ($CI_{95} = [0.2246, 0.2264]$, $t = 489.96$, $df = 26,999$, $P < 0.001$). We further found that the results were robust to drop-one analysis, showing that the effects were not driven by a single organization, and replicated without the use of bootstrapping (see *SI Appendix* for full details of these analyses).

In addition, we show that the increase in similarity between the semantic meanings of women and agency was not merely a

product of the CEO (who is now a woman) being described performing agentic actions as part of their role. The event of a woman being hired as CEO means that public organizational documents are likely to refer to her and her actions as CEO. Hence, the observed effects could be driven by organizations describing the actions of their female CEOs, rather than strategically or incidentally associating their female leaders—or women more broadly—with agentic traits. We alleviated the concern that the observed changes could be limited to descriptions of CEOs' behavior in a number of ways. First, we found that there was no negative effect on the male-agency association when female CEOs were hired [Fig. 3; average $\beta_3 = 0.033$, $SD(\beta_3) = 0.071$], which would be predicted if our results were purely driven by changes in the gender of the actor who is described performing the role of the CEO. Second, the effects replicated in direction and magnitude when controlling for the relative frequency of female references in our training documents [average $\beta_3 = 0.220$, $SD(\beta_3) = 0.077$, $t = 470.2$, $df = 26,999$, $P < 0.001$], showing that the contexts in which women are described explain our results, rather than the quantity of

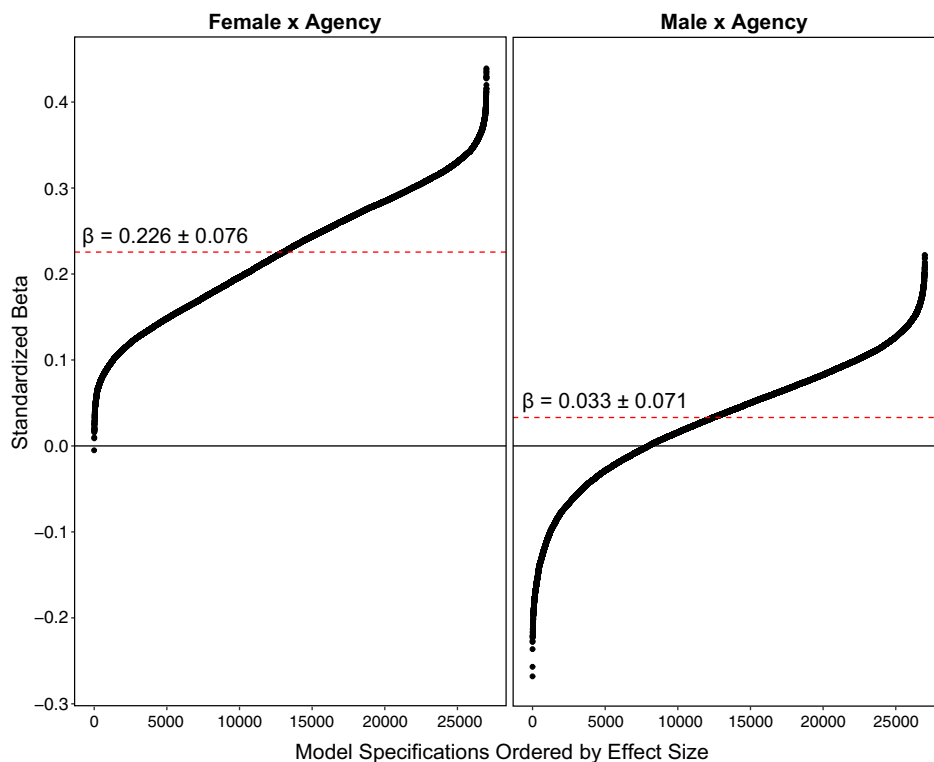


Fig. 3. Specification curve plotting the standardized beta coefficients of the interaction term β_3 from the regression Eqs. 1 and 2.

female references. Third, we found that the agentic terms driving the effect were predominantly adjectives and not action verbs (see Fig. 5). Finally, we aligned the semantic vector spaces for each organization between the pre-hiring and post-hiring periods [using Procrustes analysis (35)] and examined word-level shifts, finding that the effects were driven by a congruence of the meaning of womanhood and agency, rather than a general change in semantic meanings or instability in the concept of gender (see *SI Appendix* for more information).

Can Hiring a Female CEO Bypass the Competence–Warmth Trade-Off? Prior research has highlighted a double bind for women (5, 6). Nonagentic women are seen as warm but incompetent. In contrast, agentic women are seen as competent, but they are often characterized by the more negative aspects of agency (e.g., being dominant) and perceived as less likeable. As we have argued in the introduction, the specific nature of our intervention—hiring women into positions of leadership—is likely to mitigate backlash when it comes to the ways in which women are described in organizational documents. This is because organizations who hire women CEOs should be motivated to describe their female leaders as both positively agentic and likeable.

We test this proposition empirically using two different sets of analyses. First, we show that the shift in the female–agency association was predominantly driven by the positive aspects of agency. Second, we show that the increase in the female–agency association did not come at the expense of a decrease in the female–communality association.

Increases to Women’s Association with Agency Are Driven by the Positive Aspects of Agency. To explore whether the association of the semantic meaning of female and agency words changed equally across the positive and negative aspects of agency, we divided the agency dictionary along two relevant dimensions: 1) positive versus negative words (e.g., analytical versus aggressive)

and 2) competence-related versus dominance-related words (e.g., capability versus power) (see *SI Appendix* for the rating and validation procedure). Competence and dominance are two widely theorized facets of agency that differ in terms of their associated valence (14). While competence has clear positive connotations, dominance is oftentimes also associated with negative characteristics (14). Splitting the dictionary into sub-dictionaries allowed us to explore whether the changes to the female–agency association were driven by greater positive agency being conferred to women or backlash women faced due to organizations hiring a female leader who did not conform to people’s prototypes of leaders (6).

Fig. 4 *A–D* shows the specification curves estimated separately for each of these four dictionaries; positive, negative, competent, and dominant agency. We observed a much stronger effect ($t = 198.48$, $df = 8,999$, $P < 0.001$) for the positive facets of agency [average $\beta_3 = 0.236$, $SD(\beta_3) = 0.075$] than the negative facets [average $\beta_3 = 0.078$, $SD(\beta_3) = 0.149$]. Further, we found that both competent [e.g., intelligent, average $\beta_3 = 0.185$, $SD(\beta_3) = 0.049$] and dominant [e.g., powerful, average $\beta_3 = 0.202$, $SD(\beta_3) = 0.127$] agency became more strongly associated with women for organizations that hired female CEOs. Although the effect for dominant agency was significantly larger than the effect for competent agency ($t = 18.89$, $df = 8,999$, $P < 0.001$), the difference was negligible with regard to effect size (average β_3 of 0.202 versus 0.185).

The findings of the regression analyses were reflected in the raw cosine associations observed for the target and propensity-matched organizations pre-hire and post-hire (see tables in Fig. 4 *A–D*). Whereas the association with positive agency increased for target organizations in the period post-hire (0.0549 versus 0.0547), it decreased on average for the propensity matches (0.0544 versus 0.0548). This pattern was mirrored for both competent and dominant agency. For negative agency, the association decreased for both target and propensity matched organizations, with a steeper decrease observed among the propensity matched organizations.

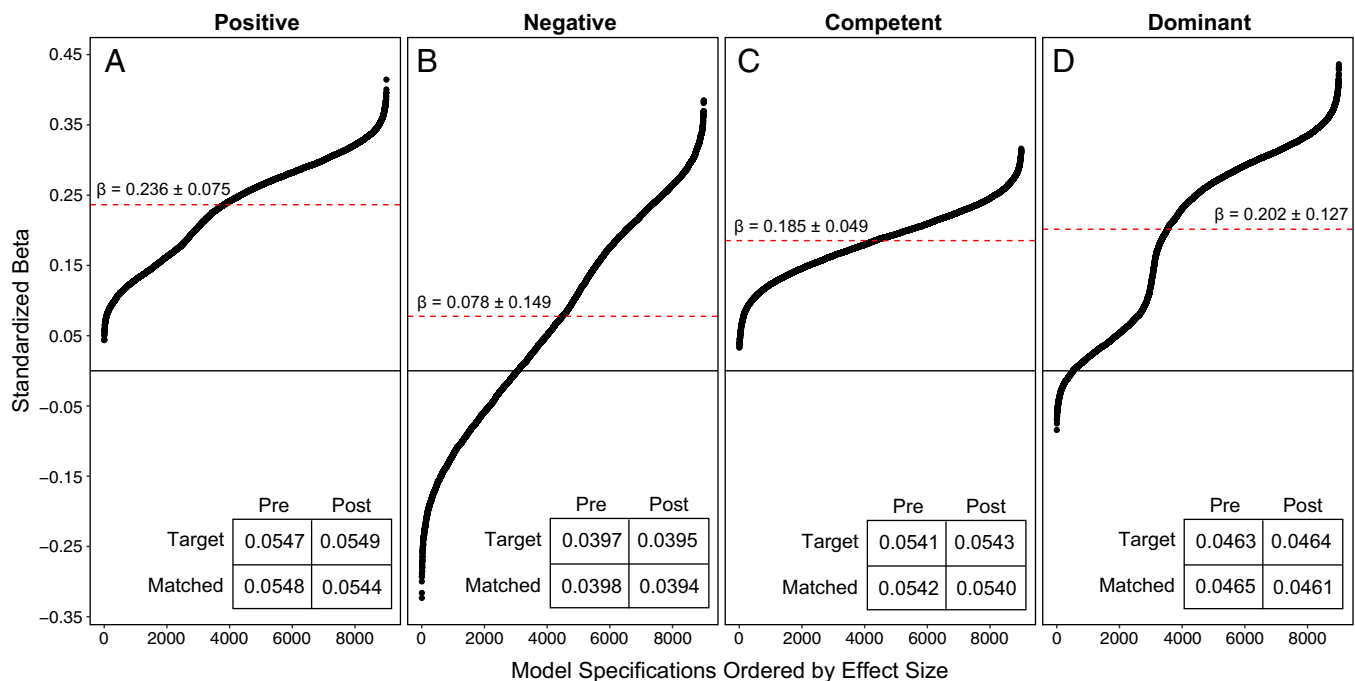


Fig. 4. Specification curves plotting the standardized beta of the interaction term β_3 from the regression Eq. 1 when splitting agency dictionary between (A and B) positive/negative and (C and D) competent/dominant.

To provide further granularity and insight into the effects, we investigated the contribution of each agentic term to the overall effect [see *SI Appendix* for further details of this procedure; (70)]. Fig. 5 displays the 30 agency words that experienced the largest positive shift in their semantic association with womanhood among the target organizations. These word-level changes are consistent with a positive account of our effects: the more negative words (e.g., dominant) are dwarfed by positive aspects of agency (e.g., direct). Notably, many of the words which saw the largest increases in their associations with the semantic meaning of female words were adjectives (e.g., active, tough,

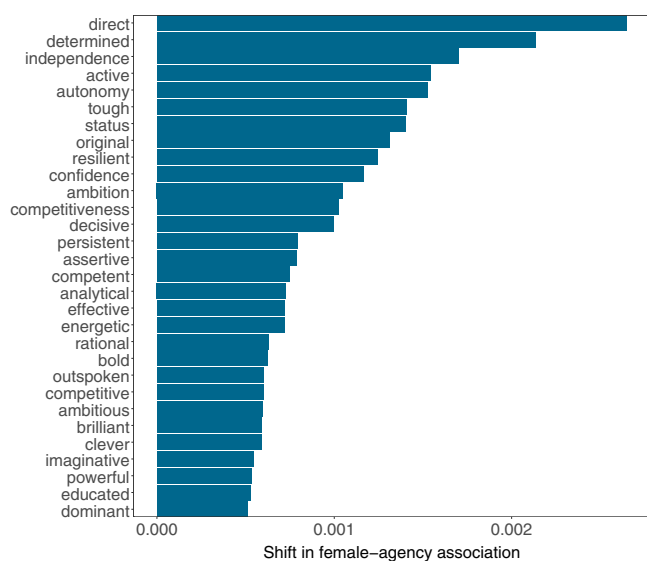


Fig. 5. The 30 agency words that saw the biggest shift in their semantic female-agency association as a function of hiring a female CEO in the target organizations.

and original), showing that our results were driven by women's association with traits, rather than behaviors.

Hiring a Female CEO Is Not Associated with a Reduction in the Female-Community Association. To understand whether the positive effect on agency comes at the cost of likeability, we studied the effect of hiring a female CEO on another gender stereotype: communality. Following the same analytical protocol as for agency, we did not observe any meaningful interaction effect on the association between female and communality related terms (Fig. 6). In fact, we observed a very small but positive effect for communality (average $\beta_3 = 0.050$, $t = 101.91$, $df = 26,999$, $P < 0.001$). Inspection of the descriptive statistics revealed that there was a small increase in the association between female and communality vectors between periods for target organizations, while there was a small decrease for the propensity matched organizations. The positive but very small effect suggests that hiring a female CEO is not associated with backlash or an agency-communality trade off in organizational language (5, 6) and that the effect of hiring a female CEO on stereotypes was specific to agency—which is role-congruent with leadership—rather than all gender stereotypes.

Overall, we found robust evidence that hiring a female CEO is associated with changes to organizational language, with female words becoming more closely related to agentic, leadership-congruent traits, without imposing a trade-off with likeability. However, it remains unclear to what extent these effects are specific to the hiring of female CEOs or can be observed when hiring women in senior leadership positions more broadly. In addition, the small sample size of only 11 target companies that hired a female CEO (and met our inclusion criteria) limits the generalizability of our effects, and the use of propensity matching implies that different results might be obtained with different matching variables. In study 2 we therefore broaden our scope to study the relationship between the proportion of women on organizational boards and the gendering of language over time across a large sample of S&P 500 companies.

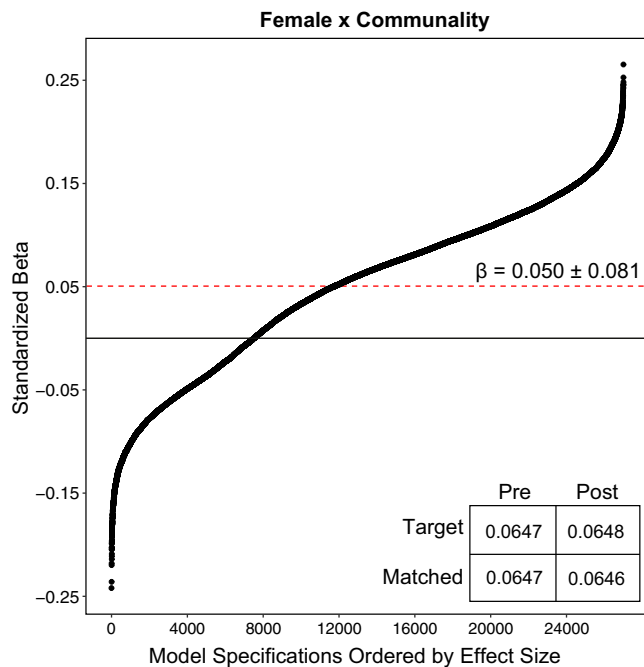


Fig. 6. Specification curves plotting the standardized beta of the interaction term β_3 from regressing the female–communality association on target, period, and target \times period.

Study 2: Female Representation on Executive Boards Is Associated with Changes in the Gendering of Language

Similar to study 1, we used word embeddings to convert organizational documents into measures of gender stereotypes (women’s association with agency and communality). However, instead of focusing on the event of hiring a female CEO, we studied the relationship between the proportion of women on companies’ executive boards and the female–agency association longitudinally across 345 former and current S&P 500 companies. This approach allowed us to replicate the findings from study 1 in a way that 1) moves past a propensity-matched quasi-experimental design, 2) can disentangle the bidirectional relationship between female representation and gendered language, 3) shows that changes to gender stereotypes generalize beyond references to the CEO, and 4) highlights the benefits of female representation in senior leadership positions more broadly.

We created a longitudinal panel in the following steps: First, we randomly selected 345 of the 690 eligible organizations that entered the S&P 500 from 2009 to 2018 to pretrain a word2vec model using the same text data as study 1. Second, we updated word2vec models for the remaining 345 organizations using a 3-y sliding window (i.e., 2009 to 2011, 2010 to 2012, etc.), providing us with eight estimates per company. Third, we estimated the average cosine similarity between gendered words and agency/communality words for each model (akin to our approach described in study 1). Finally, we merged this panel with the average proportion of women on each organization’s board of directors for the relevant period as extracted from Institutional Shareholder Services (ISS) Director Data via Wharton Research Database Services (WRDS). After removing periods for which we either had no unique text documents or board data, we were left with 2,286 observations associated with 328 unique organizations.

We estimated panel vector autoregression [PVAR (36)] models to test how the proportion of women and the gendering of language affected each other without having to specify the direction of the effects a priori (see *SI Appendix, Table S13* for

a full model output). PVAR uses a model that contains both lagged endogenous variables and unobserved individual effects, allowing us to test for causal relationships with observational data (36, 37). However, the requirements for consistent estimation of the parameters of such a model are demanding as OLS estimates are not consistent (38). PVAR solves this problem by employing an instrumental variable strategy that uses lags of endogenous variables as exogenous instruments and using a general method of moments estimator (39, 40) with an appropriate model and moment selection procedure (41).

As expected, we found a significant positive effect of the lagged proportion of women on our measure of the female–agency association ($b = 2.794$, $P = 0.039$). A 1-SD shock to the proportion of women on an organization’s board of directors (roughly equivalent to appointing one additional woman) was associated with a 0.31-SD increase to the female–agency association across the following seven periods (see *SI Appendix* for further details). Importantly, as with the appointment of women as CEOs, the increase in the female–agency association in organizational language did not come at the expense of women becoming less associated with communality. The effect of lagged proportion of women on the board of directors on the female–communality association was, in fact, positive but nonsignificant ($b = 1.517$, $P = 0.329$). Further, we also found preliminary evidence for a relationship going in the opposite direction, with the effect of the lagged female–agency association on the proportion of women being positive and statistically significant ($b = 0.006$, $P = 0.007$). While this suggests that increases to the female–agency association may also precipitate future increased female representation, this effect was much smaller than that of the proportion of women on the female–agency association. Specifically, a 1-SD shock to the female–agency association was associated with a 1–percentage point increase to the proportion of women on the board. Yet the bidirectional nature of the effect highlights the potential for a virtuous cycle: the more women get hired into leadership positions, the more embracing organizational language becomes of women’s leadership competencies, and as gender stereotypes dissipate for women more broadly, the more likely organizations are to hire additional female leaders.

Overall, the findings from study 2 provide additional evidence for the relationship between female representation in leadership positions and a reduction of gender stereotypes in organizational language. We show that this relationship generalizes to a wider range of S&P 500 organizations. While we cannot make strong claims about whether the observed changes in language reflect a general shift in how organizations think about gender, study 2 shows that the mitigation of stereotypes extends beyond female CEOs to other women on other senior leadership roles. Although the changes could still be limited to the specific women appointed, board members are much less likely to be the sole subject of organizational documents, suggesting that the semantic changes may indeed generalize to women beyond CEOs and executive board members (see *Discussion* for more details). In addition, the findings highlight the bidirectional nature of the relationship between female representation and gender stereotypes.

Discussion

The language we use to describe men and women speaks volumes and has consequences for stereotypes, career outcomes, and beyond. Here we aim to understand how semantic meaning changes as a function of appointing women as CEOs or to the board of directors, demonstrating that hiring a female CEO or board member is associated with shifts in the semantic relationship between what it means to be a woman and what it means to be a leader. Our findings contribute to the existing literature

in three important ways. Theoretically, we contribute to research on gender equality and stereotype change. Over the past few decades, multimillion-dollar efforts have been devoted to breaking down gender norms and changing what it means to be a woman: from #LikeAGirl (42) to #ShesALady (43), there has been an effortful push to change the meaning of womanhood to incorporate agentic traits like confidence, power, and independence. Such attempts are meant to change stereotypes and increase women's confidence and advancement into positions of power. Indeed, much empirical research has espoused the importance of stereotype change as a precursor to gender equality (2, 4). Yet, here we show that when it comes to stereotype change, female representation is not merely an end but also a means: simply hiring a female CEO or board member may shift the semantic association between what it means to be a woman and what it means to be a leader. Further, while previous research on backlash avoidance has demonstrated that individual behaviors (such as reframing agency as, or coupling agency with, communality) can mitigate backlash (44, 45), here we move beyond individual approaches and toward systemic ones. In our findings, increases to women's agency are driven by the positive aspects of agency and occur without corresponding costs to communality. These findings suggest that when organizations are incentivized to portray their female leaders in positive ways, stereotypes can change in a way that is both systemic, removing the burden of stereotype change from the individual, and positive, in that it does not reflect a trade-off between appearing competent and likeable.

Methodologically, we contribute to the nascent literature of using word embeddings to quantify gender stereotypes and social change (32, 33, 46). Recent designs have used word embedding models trained on large corpora of historical text to chart how the understanding of concepts has changed over time. For example, Garg et al. (33) showed that women became more or less associated with different occupations (e.g., mechanic and housekeeper) over the course of the 20th century. In contrast, our study designs aim to isolate the effect of specific events (appointing women as CEOs and to the board of directors) on the meanings imbued in word embeddings at an organization level. This approach offers broader possibilities for understanding the relationships between the semantic meanings encoded in language and real-world outcomes and provides a template for further research studying word embedding models that are associated with specific organizations, institutions, and public figures.

Practically, this work contributes to a growing call and eager push for increasing women's representation in positions of power, due to the many benefits this representation is expected to bring, both to organizations and to society. Women's participation in management has been shown to empower subordinates (47), signal change (48), lead to better management of conflict (49), and improve decision-making and performance (50, 51). However, the benefits of hiring a female leader can cascade beyond these direct benefits conferred to their organizations. Female exemplars send signals to others: they empower other women (52), affect their career aspirations (53), and improve their performance in counterstereotypic fields (54). In the present research we identify an additional benefit to increasing female representation in positions of leadership: the potential for changing pernicious gender stereotypes that are encoded in language (4). Given the many hurdles women face along their career paths, changing such stereotypes is imperative in the push toward gender equality.

There are several key limitations of the present research that should be addressed in future work. First, to what extent is the observed effect causal? Although we cannot speak definitively to causality (as we cannot randomly assign female CEOs to organizations), we attempted to support causal inference in two

ways. In study 1's quasi-experimental design, the comparison of target firms to propensity matches controlled for both time trends and relevant firm-level factors. We found no evidence of changes to the female–agency association in the prehire period (*SI Appendix, Fig. S15*), suggesting that an unobservable factor would have to cause hiring before it began to affect language. Further, study 2 employs PVAR, which is a cross-lagged panel model that aims to isolate how two or more variables are causally associated with each other over time. Research has suggested that cross-lagged panel models may be able to support causal inference by accounting for confounds and reverse causation (55). The causal interpretation of PVAR models can be controversial: in the absence of an appropriate lag testing procedure, causal inference is inappropriate (36), and if the residuals in the specification are correlated, the coefficient estimates may not be consistent (37). Although these do not appear to be problems in our analyses, future research should seek to further strengthen the case for the causality of this relationship by pursuing additional methods of analysis. For a review of PVAR models, see Canova and Ciccarelli (56).

Second, to what extent do the effects generalize beyond the appointed female leaders and impact perceptions of women more broadly? Our theoretical reasoning suggests that female representation in senior leadership positions changes gender stereotypes by providing female role models, who become associated with leadership-relevant, agentic qualities. The changes to how the appointed women are perceived are then expected to transfer to the wider category—conferring greater agency to women in general. While we observe a shift in the gendering of language when women are appointed to senior leadership positions, we cannot definitively say whether this change extends beyond the characterization of the female CEOs and board members. However, our findings provide tentative evidence that this spillover effect is likely. First, as *SI Appendix, Fig. S15* shows, the association between women and agentic characteristics continuously increases in the posthire period. If the effects were merely driven by organizational documents referencing the new female CEO, one would expect the effect to be strongest at the point of hiring. The continuous upward trend suggests that as the female leader fulfills her role, perceptions of women as being agentic may extend to women more broadly. In addition, the findings of study 2 are unlikely to solely be explained by direct references to board members, given that they are not typically discussed extensively during investor calls or in 10-K documents. However, future research should investigate this proposition more directly by trying to disentangle references to specific women leaders versus women in general, by studying the temporal dynamics of how perceptions of women change after the appointment of female leaders, and by expanding the data sources to organizational documents such as emails and minutes from meetings.

Third, at what level of seniority do the benefits of female representation cease to be present? Our findings show that appointing women as CEOs or executive board members is associated with a reduction in gender stereotypes in organizational language. However, it remains unclear whether we would expect similar effects when increasing the number of female middle managers, for example. At lower tiers of the organizational hierarchy, both the visibility of female leadership and the incentives for organizations to signal the competence (and warmth) of their female workforce might be significantly reduced compared to senior leadership positions. Thus, any changes to stereotypes may be either negligible or more ambivalent in terms of their valence and the extent to which they circumvent the typical warmth–competence trade-offs faced by women.

Fourth, to what extent do the results generalize to other organizations? Most notably, the sample of organizations that hired a female CEO in study 1 was relatively small ($n = 11$).

This could mean that idiosyncratic features of these organizations might distort our results. Importantly, the convergent evidence of study 2 largely mitigates this concern, finding the same result across 328 organizations. However, our analyses are still exclusively focused on organizations from the S&P 500, which tracks large organizations listed in US stock exchanges. The effects of hiring women to senior leadership position may be different in other cultures (57) and in smaller organizations in which leaders are more directly embedded in the organizational culture. Similarly, the effect of appointing women to senior leadership positions in other domains—such as academia or the government—may produce different effects as a function of the particular context.

Finally, our inferences were based on use of the word2vec algorithm, which has recently been shown to produce word embedding estimates that might not always be stable across situations (58, 59). In order to address this potential shortcoming, we replicated our results across different initializations, vector sizes, and hyperparameter combinations (*SI Appendix*, Figs. S4 and S5). While the robustness of our effects increases our confidence in the validity of the results, future research should apply other natural language processing algorithms such as Global Vectors for Word Representation (GloVe) (60) to corroborate our findings.

Taken together, our findings suggest that female representation in leadership positions can induce language change, such that the semantic meaning of being a woman and being agentic become more similar. While previous work had suggested that changing the gendering of language could provide a fruitful avenue for changing gender stereotypes and discrimination in the workplace, the question of how to change language remained largely unanswered. Our results suggest that hiring female leaders might not only combat discrimination by making the top seem possible for other women but might also indirectly do so by changing our understanding of what it means to be a female leader.

Methods

Dictionary Development and Validation. We jointly developed dictionaries for agency and communality based on existing literature (9, 13, 16, 61–63). First, we conducted a comprehensive literature review of papers published on agentic and communal language (see *SI Appendix* for further details). We identified 356 agency-relevant and 278 communality-relevant words and phrases. After removal of phrases and duplicate items, we retained 221 agency words and 162 communality words. To empirically establish the validity of these dictionaries and select the most relevant words, we recruited 50 independent raters affiliated with a large university on the West Coast (64% female, 36% Caucasian, 62% current students, average age 28.7 y) to rate all 383 words on multiple dimensions. To provide raters with an understanding of the constructs of agency and communality, we adapted the methods used in Abele and Wojciszke (61) and consequently acquired ratings of the agency, communality, and valence (61, 63, 64) associated with each word (see *SI Appendix* for more details).

The agency and communality dictionaries were defined as the 100 words with the highest agency and communality ratings, respectively (see *SI Appendix* for full dictionaries and interrater reliabilities). Three words were present in the top 100 ratings for both agency and communality (“hard-working,” “dedicated,” and “responsible”). These three words were removed from both dictionaries and replaced with the next highest rated words.

The gender dictionaries were derived from the female and male dictionaries of the Linguistic Inquiry Word Count software (65, 66). We used the “words that start with” function of The Free Dictionary (<https://www.thefreedictionary.com>) to manually expand words for which only a stem was specified (e.g., brother*). This resulted in an initial dictionary of 221 female words and 190 male words. To ensure that the words in the female and male dictionaries were meaningful in the context of the current analyses and differed only with regard to gender, we only retained those words that were both organizationally relevant and had an opposite-gender analog (e.g., “he” and “she”). This left us with two dictionaries of 20 words each, for women and men (see *SI Appendix* for more information and the full dictionaries).

Selection of Target and Propensity-Matched Organizations. We chose to restrict our organizations to the S&P 500, as these organizations broadly

represented the US economy and were likely to have sufficient public text data available to make meaningful inferences from word embeddings. We sampled all of the 696 organizations that entered the S&P 500 index in the period FY2009 to FY2018 and identified all of the female CEOs using a search of ISS Directors Data (71) through WRDS. This yielded a list of 41 female CEOs. We placed several restrictions on this list in order to study the effect of hiring a female CEO on the gendering of language in organizational documents. First, we required that the female CEO be preceded by a male CEO. This also excluded newly formed organizations (i.e., if a division of a larger organization was spun out and a woman took charge of the new organization). We further required 3 y of text data be available before and after the hiring date, in order to recover sufficiently reliable estimates of how the use of language had changed. This effectively limited our hiring period to between 31 December 2011 and 1 January 2016. There were 11 target organizations that fulfilled all these criteria. The full list of female CEOs and their associated organizations is included in the *SI Appendix*. In addition to these target organizations we also identified two organizations that had changed from a female CEO to a male CEO, for which we provide case analysis in the *SI Appendix*.

To provide suggestive evidence for the causality of the effect, we identified two propensity-matched organizations for each of the 11 target organizations. We used the Capital IQ Company Screening to search for other constituents of the S&P 500 index that matched each target organization based on 1) number of employees, 2) year founded, and 3) SIC (Standard Industrial Classification) codes. These propensity matches were preregistered in October 2019 (<https://osf.io/utz29/>; see *SI Appendix* for a full list of all focal organizations).

Extraction of Organizational Documents and Text. Our text sample consisted of three text sources for all organizations that had entered the S&P 500 index for at least 1 mo within our sample period (FY2009 to FY2018): 1) DEF 14A filings, 2) 10-K filings, and 3) investor calls. The form DEF 14A is filed with the SEC when a shareholder vote is required. DEF 14A documents include information intended to inform shareholders on a particular issue and as such capture more variation in natural language use than other, more technical documents. The 10-K filings are organizations’ annual reports. These constitute a highly scrutinized and clear communication between an organization and its shareholders. Finally, investor calls involve more natural communication, with high level executives answering stakeholders’ questions. This constitutes a comprehensive public source of text that helps to identify how people communicate within an organization.

We recovered the DEF 14A and 10-K filings from WRDS and scraped the investor call transcripts from Seeking Alpha. In total, we recovered 43,396 text files, of which 29,626 were transcripts from investor calls and 13,770 were SEC formal filing documents. The text files were cleaned such that all information other than text (e.g., HTML content) was removed before the analyses. We recovered text data relating to 696 unique organizations. A total of 699 unique CIKs (Central Index Keys) were found in the SEC documents, and 690 tickers were found in the Seeking Alpha transcripts. Of the 699 CIKs and 690 tickers, 687 matched one to one, 6 CIKs had no tickers, and 6 CIKs matched two tickers each (and were listed as 3 organizations), forming the total of 696 unique organizations. For study 2, we only used the 690 organizations for whom we had matches in both the SEC and transcript data, to avoid biasing our results due to imbalanced text sources.

Estimation of Semantic Associations between Vectors. In line with the canon of NLP, we assume that the dimensions of word vectors embody semantic meaning (27, 31–34, 67). This implies that the similarity between two vectors reflects how closely associated the two words represented by these vectors are, in terms of meaning (67). To calculate our measure of the association between women and agency we calculated the cosine similarity between the vectors representing each of the words in the gender and the theory dictionaries. For example, to obtain a single measure of the association between women and agency, we averaged these similarity measures across all word dyads in our dictionaries (i.e., when combining 20 different gender vectors with 100 agency vectors, our measure was an average of 2,000 cosine similarities between each pair of vectors). The resulting female–agency association measure captures the association between women and agency in the language of a particular organization in a particular time period (prehire or posthire). These data are publicly available (69).

Study 2. We recovered director information from the ISS database in WRDS and aggregated it to recover organizational-level data regarding the gender diversity of corporate boards of directors. These data capture the number of women on each organization’s board of directors and the total number of people on the board of directors. This allowed us to express each board’s female representation as a single value—the proportion of female directors.

We averaged these yearly data across the 3 y periods that corresponded to the word2vec models and merged the two datasets. For example, Apple had a board of directors with an average of 21.7% women from 2014 to 2016, increasing to 26.1% in the period 2015 to 2017: for each of these periods, we also had a measure of the female–agency association conveyed in organizational documents. To focus on the core effect, we used 100-dimensional word vectors and the average association between the 10-word gender dictionaries and 100-word theory dictionaries to construct our measures.

PVAR (36, 68) requires the specification of many parameters. We used a generalized method of moments estimator without system instruments, a single lag of the endogenous variables, forward orthogonal transformation of lags, collapsed first difference moment conditions, and tested the robustness of our results across different specifications of which variables were endogenous. We delineate the rationale and statistical tests (where appropriate) behind each of these choices in full in the *SI Appendix*.

- Catalyst, Quick take: Women in management. *Catalyst* (13 October 2020). <https://www.catalyst.org/research/women-in-management/>. Accessed 13 August 2020.
- M. E. Heilman, Description and prescription: How gender stereotypes prevent women's ascent up the organizational ladder. *J. Soc. Issues* **57**, 657–674 (2001).
- A. H. Eagly, S. J. Karau, Role congruity theory of prejudice toward female leaders. *Psychol. Rev.* **109**, 573–598 (2002).
- N. Ellemers, Gender stereotypes. *Annu. Rev. Psychol.* **69**, 275–298 (2018).
- L. A. Rudman, C. A. Moss-Racusin, P. Glick, J. E. Phelan, "Reactions to vanguards: Advances in backlash theory" in *Advances in Experimental Social Psychology*, P. Devine, A. Plant, Eds. (Academic Press, 2012), chap. 4, pp. 167–227.
- L. A. Rudman, J. E. Phelan, Backlash effects for disconfirming gender stereotypes in organizations. *Res. Organ. Behav.* **28**, 61–79 (2008).
- E. H. Chang *et al.*, The mixed effects of online diversity training. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 7778–7783 (2019).
- P. G. Devine *et al.*, A gender bias habit-breaking intervention led to increased hiring of female faculty in STEM departments. *J. Exp. Soc. Psychol.* **73**, 211–215 (2017).
- A. E. Martin, K. W. Phillips, What "blindness" to gender differences helps women see and do: Implications for confidence, agency, and action in male-dominated environments. *Organ. Behav. Hum. Decis. Process.* **142**, 28–44 (2017).
- K. Kawakami, J. F. Dovidio, S. van Kamp, Kicking the habit: Effects of nonstereotypic association training and correction processes on hiring decisions. *J. Exp. Soc. Psychol.* **41**, 68–75 (2005).
- M. J. Zawadzki, S. A. Shields, C. L. Danube, J. K. Swim, Reducing the endorsement of sexism using experiential learning: The Workshop Activity for Gender Equity Simulation (WAGES). *Psychol. Women Q.* **38**, 75–92 (2014).
- S. Kesebir, Word order denotes relevance differences: The case of conjoined phrases with lexical gender. *J. Pers. Soc. Psychol.* **113**, 262–279 (2017).
- J. M. Madera, M. R. Hebl, R. C. Martin, Gender and letters of recommendation for academia: Agentic and communal differences. *J. Appl. Psychol.* **94**, 1591–1599 (2009).
- A. S. Rosette, C. Z. Koval, A. Ma, R. Livingston, Race matters for women leaders: Intersectional effects on agentic deficiencies and penalties. *Leadersh. Q.* **27**, 429–445 (2016).
- M. Olsson, S. E. Martiny, Does exposure to counterstereotypical role models influence girls' and women's gender stereotypes and career choices? A review of social psychological research. *Front. Psychol.* **9**, 2264 (2018).
- S. Correll, K. Weishaar, K. A. Wynn, J. Wehner, Inside the black box of organizational life: The Gendered language of performance assessment. *Am. Sociol. Rev.* **85**, 1022–1050 (2020).
- A. Marchant, A. Bhattacharya, M. Carnes, Can the language of tenure criteria influence women's academic advancement? *J. Womens Health (Larchmt.)* **16**, 998–1003 (2007).
- L. Boroditsky, L. A. Schmidt, W. Phillips, "Sex, syntax, and semantics" in *Language in mind: Advances in the study of language and thought*, D. Gentner and S. Goldin-Meadow, Eds. (MIT Press, Cambridge, MA, 2003), pp. 61–79.
- S. Sczesny, M. Formanowicz, F. Moser, Can gender-fair language reduce gender stereotyping and discrimination? *Front. Psychol.* **7**, 25 (2016).
- J. E. Phelan, L. A. Rudman, Prejudice toward female leaders: Backlash effects and women's impression management dilemma. *Soc. Personal. Psychol. Compass* **4**, 807–820 (2010).
- N. Dasgupta, S. Asgari, Seeing is believing: Exposure to counterstereotypic women leaders and its effect on the malleability of automatic gender stereotyping. *J. Exp. Soc. Psychol.* **40**, 642–658 (2004).
- E. R. Smith, M. A. Zarate, Exemplar-based model of social judgment. *Psychol. Rev.* **99**, 3–21 (1992).
- B. L. Connelly, S. T. Certo, R. D. Ireland, C. R. Reutzel, Signaling theory: A review and assessment. *J. Manage.* **37**, 39–67 (2011).
- S. T. Certo, Influencing initial public offering investors with prestige: Signaling with board structures. *Acad. Manage. Rev.* **28**, 432–446 (2003).
- V. C. Burbano, Social responsibility messages and worker wage requirements: Field experimental evidence from online labor marketplaces. *Organ. Sci.* **27**, 1010–1028 (2016).
- D. A. Jones, C. R. Willness, S. Madey, Why are job seekers attracted by corporate social performance? Experimental and field tests of three signal-based mechanisms. *Acad. Manage. J.* **57**, 383–404 (2014).
- T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, J. Dean, "Distributed representations of words and phrases and their compositionality" in *Advances in Neural Information Processing Systems* **26**, C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, K. Q. Weinberger, Eds. (Curran Associates, Inc., 2013), pp. 3111–3119.
- Y. Goldberg, O. Levy, word2vec explained: Deriving Mikolov et al.'s negative-sampling word-embedding method. arXiv [Preprint] (2014). <https://arxiv.org/abs/1402.3722> (Accessed 17 April 2020).
- X. Rong, word2vec parameter learning explained. arXiv [Preprint] (2016). <https://arxiv.org/abs/1411.2738> (Accessed 17 April 2020).
- J. R. Firth, A Synopsis of Linguistic Theory: Studies in Linguistic Analysis (Blackwell, Oxford, UK, 1957).
- A. Caliskan, J. J. Bryson, A. Narayanan, Semantics derived automatically from language corpora contain human-like biases. *Science* **356**, 183–186 (2017).
- D. DeFranza, H. Mishra, A. Mishra, How language shapes prejudice against women: An examination across 45 world languages. *J. Pers. Soc. Psychol.* **119**, 7–22 (2020).
- N. Garg, L. Schiebinger, D. Jurafsky, J. Zou, Word embeddings quantify 100 years of gender and ethnic stereotypes. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E3635–E3644 (2018).
- T. Bolukbasi, K.-W. Chang, J. Y. Zou, V. Saligrama, A. T. Kalai, "Man is to computer programmer as woman is to homemaker? Debiasing word embeddings" in *Advances in Neural Information Processing Systems* **29**, D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, R. Garnett, Eds. (Curran Associates, Inc., 2016), pp. 4349–4357.
- J. C. Gower, Generalized procrustes analysis. *Psychometrika* **40**, 33–51 (1975).
- D. Holtz-Eakin, W. Newey, H. S. Rosen, Estimating vector autoregressions with panel data. *Econometrica* **56**, 1371–1395 (1988).
- J. D. Angrist, J.-S. Pischke, *Mostly Harmless Econometrics: An Empiricist's Companion* (Princeton University Press, 2009).
- S. Nickell, Biases in dynamic models with fixed effects. *Econometrica* **49**, 1417–1426 (1981).
- M. Arellano, S. Bond, Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Rev. Econ. Stud.* **58**, 277–297 (1991).
- R. Blundell, S. Bond, Initial conditions and moment restrictions in dynamic panel data models. *J. Econ.* **87**, 115–143 (1998).
- D. W. K. Andrews, B. Lu, Consistent model and moment selection procedures for GMM estimation with application to dynamic panel data models. *J. Econ.* **101**, 123–164 (2001).
- M. Lee, H. J. Yoon, When brand activism advertising campaign goes viral: An analysis of always #LikeAGirl video networks on YouTube. *Int. J. Adv. Cult. Technol.* **8**, 146–158 (2020).
- N. Gillil, Why women are talking about H&M's latest ad campaign. *Econsultancy* (2016). <https://econsultancy.com/why-women-are-talking-about-h-m-s-latest-ad-campaign/>. Accessed 5 May 2021.
- M. E. Heilman, T. G. Okimoto, Why are women penalized for success at male tasks?: The implied communality deficit. *J. Appl. Psychol.* **92**, 81–92 (2007).
- L. A. Rudman, P. Glick, Prescriptive gender stereotypes and backlash toward agentic women. *J. Soc. Issues* **57**, 743–762 (2001).
- A. C. Kozlowski, M. Taddy, J. A. Evans, The geometry of culture: Analyzing the meanings of class through word embeddings. *Am. Sociol. Rev.* **84**, 905–949 (2019).
- B. M. Bass, B. J. Avolio, L. Atwater, The transformational and transactional leadership of men and women. *Appl. Psychol.* **45**, 5–34 (1996).
- M. K. Ryan, S. A. Haslam, M. D. Hershey, R. Bongiorno, Think crisis-think female: The glass cliff and contextual variation in the think manager-think male stereotype. *J. Appl. Psychol.* **96**, 470–484 (2011).
- S. E. Perkins, K. W. Phillips, N. A. Pearce, Ethnic diversity, gender, and national leaders. *J. Int. Aff.* **67**, 85–104 (2013).
- C. L. Dezsó, D. G. Ross, Does female representation in top management improve firm performance? A panel data investigation. *Strateg. Manage. J.* **33**, 1072–1089 (2012).
- K. Sergeant, A. D. Stajkovic, Women's leadership is associated with fewer deaths during the COVID-19 crisis: Quantitative and qualitative analyses of United States governors. *J. Appl. Psychol.* **105**, 771–783 (2020).

52. C. Wolbrecht, D. E. Campbell, Leading by example: Female members of parliament as political role models. *Am. J. Pol. Sci.* **51**, 921–939 (2007).
53. D. M. Young, L. A. Rudman, H. M. Buettner, M. C. McLean, The influence of female role models on women's implicit science cognitions. *Psychol. Women Q.* **37**, 283–292 (2013).
54. S. D. Herrmann *et al.*, The effects of a female role model on academic performance and persistence of women in STEM courses. *Basic Appl. Soc. Psychol.* **38**, 258–268 (2016).
55. M. J. Zyphur *et al.*, From data to causes I: Building a general cross-lagged panel model (GCLM). *Organ. Res. Methods* **23**, 651–687 (2020).
56. F. Canova, M. Ciccarelli, "Panel vector autoregressive models: A survey" in *Advances in Econometrics*, vol. 32, T. B. Fomby, L. Kilian, A. Murphy, Eds. (Emerald Group Publishing Limited, 2013), pp. 205–246.
57. A. J. C. Cuddy *et al.*, Men as cultural ideals: Cultural values moderate gender stereotype content. *J. Pers. Soc. Psychol.* **109**, 622–635 (2015).
58. M. Antoniak, D. Mimno, Evaluating the stability of embedding-based word similarities. *Trans. Assoc. Comput. Linguist.* **6**, 107–119 (2018).
59. L. Wendlandt, J. K. Kummerfeld, R. Mihalcea, "Factors influencing the surprising instability of word embeddings" in *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (Association for Computational Linguistics, New Orleans, LA, 2018), vol. 1, pp. 2092–2102.
60. J. Pennington, R. Socher, C. Manning, "Glove: Global vectors for word representation" in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (Association for Computational Linguistics, 2014), pp. 1532–1543.
61. A. E. Abele, B. Wojciszke, Agency and communion from the perspective of self versus others. *J. Pers. Soc. Psychol.* **93**, 751–763 (2007).
62. S. Bruckmüller, A. E. Abele, The density of the big two. *Soc. Psychol.* **44**, 63–74 (2013).
63. M. Diehl, S. K. Owen, L. M. Youngblade, Agency and communion attributes in adults' spontaneous self-representations. *Int. J. Behav. Dev.* **28**, 1–15 (2004).
64. A. E. Martin, K. W. Phillips, Blind to bias: The benefits of gender-blindness for STEM stereotyping. *J. Exp. Soc. Psychol.* **82**, 294–306 (2019).
65. J. W. Pennebaker, R. L. Boyd, K. Jordan, K. Blackburn, *The Development and Psychometric Properties of LIWC2015* (University of Texas at Austin, Austin, TX, 2015).
66. Y. R. Tausczik, J. W. Pennebaker, The psychological meaning of words: LIWC and computerized text analysis methods. *J. Lang. Soc. Psychol.* **29**, 24–54 (2010).
67. S. Bhatia, Associative judgment and vector space semantics. *Psychol. Rev.* **124**, 1–20 (2017).
68. M. Sigmund, R. Ferstl, Panel vector autoregression in R with the package panelvar. *Q. Rev. Econ. Finance* **80**, 693–720 (2019).
69. M. A. Lawson, A. E. Martin, I. Huda, S. C. Matz, Vector Distance Data. OSF. <https://osf.io/7t95y/>. Deposited 24 November 2020.
70. M. A. Lawson, A. E. Martin, I. Huda, S. C. Matz, Vector Distance Data at Word Level. OSF. <https://osf.io/9ztp8/>. Deposited 24 November 2020.
71. Institutional Shareholder Services, Directors Data. <https://www.issgovernance.com/esg/governance-data/director-data/>. Accessed 30 November 2018.