



A Crisis of Competence: Benevolent Sexism Affects Evaluations of Women's Competence

Brittany S. Cassidy¹  · Anne C. Krendl²

© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract

People higher in benevolent sexism often outwardly endorse gender equality, but support men over women for challenging positions and experiences. Reflecting shifting standards (a tendency to evaluate stereotyped group members against within-category judgment standards), people higher in sexism may evaluate prominent women's competence against a lower competency standard for women (who are stereotyped as less competent than men are), and not against a standard for men. Thus prominent women could be perceived as especially competent (versus other women), yet men might still garner ultimate support. Study 1 tested for this possibility using an ecologically valid example: the 2016 U.S. presidential election. Study 1 showed that benevolent (and not hostile) sexism predicted less opposition to Donald Trump's candidacy and more positive attitudes toward the election outcome among 57 mostly female U.S. college students. Study 1 also showed that benevolent sexism positively predicted competence perceived in Hillary Clinton. To determine if this positive relationship reflected shifting standards, we manipulated the gender to which a prominent woman would be compared in Study 2 with 189 U.S. adults. Reflecting shifting standards, benevolent sexism related to evaluating women as more competent when they were evaluated against other women versus other men. Shifting standards also mediated a relationship between benevolent sexism and expecting lower female success. Using shifting standards may be one way that people higher in benevolent sexism might evaluate prominent women as especially competent, yet ultimately support men.

Keywords Sexism · Competence · Stereotyping · Impressions · Shifting standards

Despite recent declines in overt displays of sexism (Dovidio and Gaertner 1998), sexism continues to negatively affect women (Glick et al. 2000) seeking prominent leadership positions both in the workplace (Hideg and Ferris 2016) and on the political stage (Gervais and Hillard 2011). The negative consequences of sexism undermine a societal ideal of gender equality. Indeed, approximately 40% of U.S. Financial Post 500 companies have no women on their board of directors

(Catalyst 2013). Because perceived competence is related to career success (e.g., Todorov et al. 2005), understanding how sexism relates to perceptions of competence in prominent women remains a critical area of research.

Ambivalent sexism theory has illustrated two correlated but distinct aspects of sexism (Glick and Fiske 1996, 1997). Hostile sexism reflects a conceptualization of sexism whereby traditional gender roles are maintained through derogatory characterizations of women. By contrast, benevolent sexism is subtler (Glick and Fiske 2001) and maintains these roles using more positive characterizations (e.g., women need men's help). Despite seemingly more positive characterizations, benevolent sexism assumes that women comprise the weaker, and thereby less competent, sex. Because benevolent sexism is associated with positive characterizations of women, but discriminatory behavior toward them, the current investigation tested for the possibility that benevolent sexism could inform the discrepancy between perceptions that women are competent to perform a challenging job (e.g., hold political office), but lack of support for women having those jobs.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s11199-019-1011-3>) contains supplementary material, which is available to authorized users.

✉ Brittany S. Cassidy
bscassid@uncg.edu

¹ Department of Psychology, University of North Carolina at Greensboro, 296 Eberhart, PO Box 26170, Greensboro, NC 27412, USA

² Department of Psychological and Brain Sciences, Indiana University, Bloomington, 1101 E. 10th St., Bloomington, IN 47405, USA

Benevolent sexism contributes to gender inequality by increasing support for traditional gender roles (i.e., that men belong in leadership positions; Jost and Kay 2005). Indeed, benevolent sexism is related to fewer women involved in the economy and in politics (Glick et al. 2000). People higher in benevolent sexism are less likely to be perceived as openly sexist (Barreto and Ellemers 2005) despite behaviors that perpetuate gender inequality (Becker and Wright 2011). Illustrating this point and suggesting that they might outwardly evaluate men and women as similarly competent, people higher in benevolent sexism voice support for gender-based employment equality. Yet, the same people contribute to inequality by ultimately promoting men versus women (Hideg and Ferris 2016), and by assigning men (versus women) to challenging workplace experiences, even when men and women are similarly interested in them (King et al. 2012). These behaviors suggest a discrepancy in how people scoring higher in benevolent sexism evaluate women's competence and how they ultimately support them. An important, yet unexplored, question regards how higher benevolent sexism relates to perceived competence in women who have already shown competency in their professions. That is, how do more benevolently sexist people reconcile that a woman is highly competent with ultimately supporting a man? This question is important because how benevolent sexism relates to evaluations of competence in prominent women is critical for developing strategies to reduce the negative consequences of benevolent sexism for such women, such as their not being promoted. Here, we examined if shifting standards represents one mechanism by which benevolent sexism may influence competence evaluations of prominent women.

People higher in benevolent sexism may reconcile the discrepancy between the achievements of prominent women and upholding a status quo of men in power by evaluating women's competence against a different baseline. This possibility reflects shifting standards (Biernat et al. 1991) and would allow for the attitudes and behaviors of benevolent sexists to be inconsistent (for reviews, see Ajzen and Fishbein 1977, 2000). The *shifting standards* model posits that beliefs affect how individuals perceive traits by creating different evaluative standards (e.g., an expected level of competence) against which group members (e.g., women) are evaluated (Biernat et al. 1991). The key insight from this model is that people evaluate targets against a within-group standard (e.g., women only) versus a standard encompassing several groups (e.g., men and women). With respect to competence evaluations, the fact that women are stereotyped as being less competent than men are (Broverman et al. 1972) could manifest as people with higher benevolent sexism using a different (lower) reference point for evaluating women's competence (e.g., Biernat and Kobrynowicz 1997). In other

words, more prejudiced people would have a lower baseline against which to evaluate a woman's competence. Indeed, people with more prejudice are more likely to activate (Wittenbrink et al. 1997) and apply (Lepore and Brown 1997) stereotypes in their evaluations, impacting the evaluative standard they use to assess stereotyped group members (Biernat and Manis 1994).

The shifting standards model may shed light on the discrepancy between highly benevolently sexists' perceptions of a woman's perceived competence and their reluctance to support her. More strongly benevolent sexists might evaluate prominent women as having especially high competence because they are judged relative to other women, and not relative to men *and* women. At the same time, because higher benevolent sexists view women as being less competent than are men in general, they would ultimately support a man, and not a woman, for a prominent position. These patterns would further the literature by elucidating shifting standards as a potential way for people with more prejudice to evaluate some underrepresented category members positively while preserving a tradition of other category members having higher status.

Over two studies, we examined how benevolent sexism related to evaluating competence in prominent women. In Study 1, we tested whether benevolent sexism positively related to perceived competence using an ecologically valid example: Hillary Clinton. Hillary Clinton was a prominent United States politician who was generally perceived as being highly competent (Gaffney and Blaylock 2010) and conscientious (Visser et al. 2017), in part due to her more masculine leadership and communication style (Carlin and Winfrey 2009; Huddy and Terkildsen 1993). If evaluated based on her gender, benevolent sexism may *positively* correspond to trait perceptions of Clinton's competence because her achievements (Carlin and Winfrey 2009) reflect high competency relative to other women for whom a stereotype of lower competence would be more applicable. If true, Clinton could have been perceived by people higher in benevolent sexism as being very competent because she was not evaluated against her male opponents. However, the same people would be expected to ultimately support a male candidate (i.e., Donald Trump, Clinton's political opponent) and react positively toward Clinton's loss in the 2016 U.S. presidential election because benevolent sexism is associated with upholding traditional gender roles (Glick and Fiske 2001). Because the U.S. presidency has both historically and recently been perceived as a masculine position more appropriate for men than for women (Paul and Smith 2008; Rosenwasser and Dean 1989; Smith et al. 2007), ultimately supporting a male candidate would maintain those roles. In Study 2, we tested if, reflecting shifting standards, evaluating a prominent woman (a

candidate for U.S. senator) as especially competent when compared to other women versus other men positively related to benevolent sexism.

Study 1

The goals of Study 1 were threefold. Benevolent sexism is associated with maintaining traditional gender roles (Glick and Fiske 2001). We thus first sought to test if benevolent, but not hostile, sexism predicted support for men's candidacies for traditionally masculine positions (as in Hideg and Ferris 2016). We expected that benevolent sexism would predict less opposition to a man's (i.e., Donald Trump's) presidential candidacy and more positive attitudes toward the 2016 United States election outcome in which Trump won (Hypothesis 1) because such patterns would maintain traditional gender roles of men being in prominent positions—a key component of benevolent sexism. If endorsing a man's candidacy reflected distinct attitudes toward a *candidate*, however, hostile sexism would be expected to predict less opposition to Donald Trump's candidacy (similar to Ratliff et al. 2017). This could be because language used by Donald Trump during his campaign often devalued specific women (Darweesh and Abdullah 2016)—a key component of hostile sexism. If, as predicted, benevolent sexism predicted less opposition to Donald Trump's candidacy, we could then expand on this finding by examining if benevolent sexism nevertheless related to more positive characterizations of women as predicted by the literature (e.g., Glick and Fiske 2001).

Building on Hypothesis 1, we expected that benevolent sexism would positively relate to trait perceptions of Hillary Clinton's competence (Hypothesis 2). Testing for a positive relationship would lay groundwork on which to test if, supporting shifting standards, benevolent sexism predicted higher evaluations of prominent women's competence when comparing to other women versus other men. That is, a prominent woman's achievements may defy a stereotype of lower female competency held by more prejudiced individuals and allow her to be perceived as especially competent. Benevolent sexism should not positively predict perceptions of Donald Trump's competence because, as a man, he would already be expected to be competent and not defy that stereotype.

Clinton's loss allowed for the opportunity to build on Hypothesis 2 by probing the expected positive relationship between benevolent sexism and her perceived competence. Clinton's perceived competence may not have defied a lower competency standard for women held by people higher in benevolent sexism (see Biernat and Manis 1994) after she lost the election (e.g., Thoroughgood et al. 2013). This possibility means that when Clinton is perceived in a context where she

does not defy lower competency (e.g., when she lost versus when she was widely projected to win), strongly benevolent sexists should not perceive her as especially competent. If an expected positive relationship between benevolent sexism and Clinton's perceived competence reflects shifting standards, benevolent sexism should more strongly relate to Clinton's perceived competency before versus after the election (Hypothesis 3). To this end, we obtained competence perceptions of Clinton and Trump pre- and post-election.

Although Hillary Clinton exemplifies a woman who has shown high competency, she has been a polarizing political figure for decades (Carlin and Winfrey 2009). Given this polarization, it may be difficult for perceivers to provide nuanced ratings of her competence. Indeed, partisan beliefs are strongly associated with biased explicit trait ratings of candidates (Wright and Tomlinson 2018). Moreover, because the two timepoints at which data collection occurred were particularly tumultuous (before and immediately after the election), we circumvented participants explicitly evaluating Clinton's and Trump's competence by characterizing competence using reverse correlation (Mangini and Biederman 2004).

Reverse correlation is a data-driven method estimating how people visually represent traits in faces without their explicit endorsements of those traits (see Method for details; Dotsch and Todorov 2012). Critically, there is a growing body of evidence suggesting that reverse correlation provides visual reflections of how people perceive traits in others' faces (for a review, see Brinkman et al. 2017). Allowing for the possibility that benevolent sexism may influence trait perceptions of Clinton's competence, attitudes affect how traits are visually represented in faces (Dotsch et al. 2008). For example, support for U.S. presidential candidate Mitt Romney in 2012 related to more positive trait perceptions of his face (Young et al. 2014). Reverse correlation is thus nicely suited to test for a positive relationship between benevolent sexism and trait perceptions of Clinton's competence.

Because benevolent (versus hostile) sexism is characterized by positive characterizations of women, the goal of Study 1 was to test if it positively related to Clinton's perceived competence, yet also positively related to endorsing a candidacy that upheld traditional gender roles (i.e., a man's). Such a relationship would elucidate a real-world situation in which benevolent sexism negatively affects women. However, this possibility does not mean that hostile sexism might not also positively relate to Clinton's perceived competence. People higher in prejudice use shifting standards more when evaluating members of underrepresented groups (Biernat and Manis 1994), suggesting that hostile sexism might also positively relate to Clinton's perceived competence. However, hostile sexism might negatively relate to Clinton's perceived competence because it is related to more

explicit derogation of women. Because these relationships could be of interest to future research, we included exploratory analyses on how hostile sexism affects Clinton's perceived competence in our [online supplement](#).

Method

Participants

Fifty-seven U.S. Indiana University students ($M_{\text{age}} = 18.70$ years, $SD = 1.24$, range = 17–24, 46 female) from the United States provided written informed consent to participate as part of a larger study on attitudes toward the 2016 United States presidential election (held on November 8, 2016). They were compensated with course credit and \$10, respectively, for pre-election (8/31–11/4/16) and post-election (11/9–12/5/16) visits. Power analyses (G*Power; Faul et al. 2007) (using $r = .40$ and $\alpha = .05$) targeted 46 participants for 80% power to detect a relationship between benevolent sexism and opposition to Clinton's candidacy. The effect size was selected based on work showing benevolent sexism to be related to inaction in promoting women in underrepresented positions (Hideg and Ferris 2016). We continued data collection past 46 participants pre-election to ensure a sufficient sample of participants who completed the pre- and post-election visits. All 57 participants completed both visits. The Indiana University IRB approved all studies before the start of data collection.

Pre-Election Visit: Measuring Sexism and Candidate Opposition

To estimate pre-election trait competence perceptions of Clinton and Trump, participants completed two face classification tasks (described in the following) and several questionnaires, including the Ambivalent Sexism Inventory (ASI; Glick and Fiske 1996). The ASI measures hostile (e.g., "Women are too easily offended") and benevolent (e.g., "Women should be cherished and protected by men") components of sexism. Participants rated each of 11 items measuring hostile sexism and each of 11 items measuring benevolent sexism on a 6-point scale ranging from 0 (*disagree strongly*) to 5 (*agree strongly*). Responses across items were averaged, with higher scores indicating more sexism. Alphas of .74 (hostile sexism) and .76 (benevolent sexism) from recent work (Husnu 2016) demonstrate strong reliability of the ASI. The reliability of the hostile (Cronbach's $\alpha = .89$) and benevolent (Cronbach's $\alpha = .85$) sexism components in the current study were similarly reliable. The development of the ASI established its convergent, discriminant, and predictive validity over six studies and 2250 respondents (Glick and Fiske 1996).

Participants reported more benevolent ($M = 2.16$, $SD = .94$) than hostile ($M = 1.82$, $SD = .96$) sexism, $t(56) = 2.83$, $p = .006$, $d = .39$. Male and female participants did not differ in their hostile ($M_{\text{male}} = 1.47$, $SD = 1.05$; $M_{\text{female}} = 1.91$, $SD = .93$), $t(55) = 1.37$, $p = .18$, $d = .44$, or benevolent ($M_{\text{male}} = 1.83$, $SD = 1.08$; $M_{\text{female}} = 2.24$, $SD = .89$), $t(55) = 1.31$, $p = .20$, $d = .41$, sexism. Because sexism did not differ by gender, we did not include participants' gender in the analyses. Consistent with the literature (Glick and Fiske 1997), hostile and benevolent sexism were positively correlated, $r(55) = .57$, $p < .001$. Despite the fact that hostile and benevolent sexism are overlapping constructs, however, they did not perfectly map onto each other. Participants also used two 10-point scales to indicate the degree of their opposition to the candidacies of Clinton and Trump, on a scale from 1 (*not at all*) to 10 (*extremely oppose*).

Post-Election Visit: Election Attitudes

To measure post-election competence perceptions, participants completed the same two face classification tasks (described in the following). Fifty-six of 57 participants indicated their supported candidate in the election: Clinton (29), Trump (19), and other (8). Participants also indicated their opposition toward the candidacies of Clinton and Trump using the aforementioned scale, and their attitudes (four positive: satisfied, happy, proud, relieved; three negative: concerned, disappointed, and upset) toward the election outcome using 10-point scales from 1 (*not at all*) to 10 (*extremely*). We generated a composite election attitude score by averaging the four positive and three reverse coded negative attitudes (Cronbach's $\alpha = .85$). Higher scores indicate more positive attitudes toward the election outcome.

Estimating Visualizations of candidate's Faces

We used reverse correlation, a data-driven method to model internal representations of faces (Dotsch and Todorov 2012), to assess participants' trait perceptions of competence in Hillary Clinton and Donald Trump. The reverse correlation method has two phases. The first phase is a face classification phase completed by participants. Here, the face classification phase was used to estimate participants' mental representations of Clinton's face and Trump's face. The second phase is a face ratings phase completed by naïve raters unaware of how the faces they are rating were generated. Here, naïve raters rated the faces generated from the first phase (classification) on how competent they appeared. Ratings from the second phase (ratings) thus permitted us to objectively test the relationship between participants' mental

representations of Clinton's competence with their benevolent sexism without participants' explicit reports.

Face Classification Tasks

Participants first completed two face classification tasks in a counterbalanced order: one for Clinton's face and the other for Trump's face.

Stimuli Stimuli for the face classification tasks were generated from two base images: front-facing and neutrally expressive gray-scale 512 × 512 pixel images of the faces of Clinton and Trump from www.cnn.com (see Fig. 1a). Randomly generated noise patterns were generated in accordance with past work and layered over each image using the *rcicr* package for R (Fig. 1b; for complete details on the noise patterns, see Dotsch and Todorov 2012; Mangini and Biederman 2004). An image layered with a unique noise pattern and an image layered with that pattern's inverse was generated for each of 100 trials, totaling 200 images per task. The same noise patterns were used for all participants.

Face Classification Tasks Each face classification task consisted of 100 trials (e.g., Hehman et al. 2015) in which two images were presented side-by-side for 3 s. Depending on the task, participants selected the image most resembling either Clinton or Trump on each trial. Images were presented for 3 s to ensure that participants would not ruminate on their choices (as in Mangini and Biederman 2004). A benefit of the face

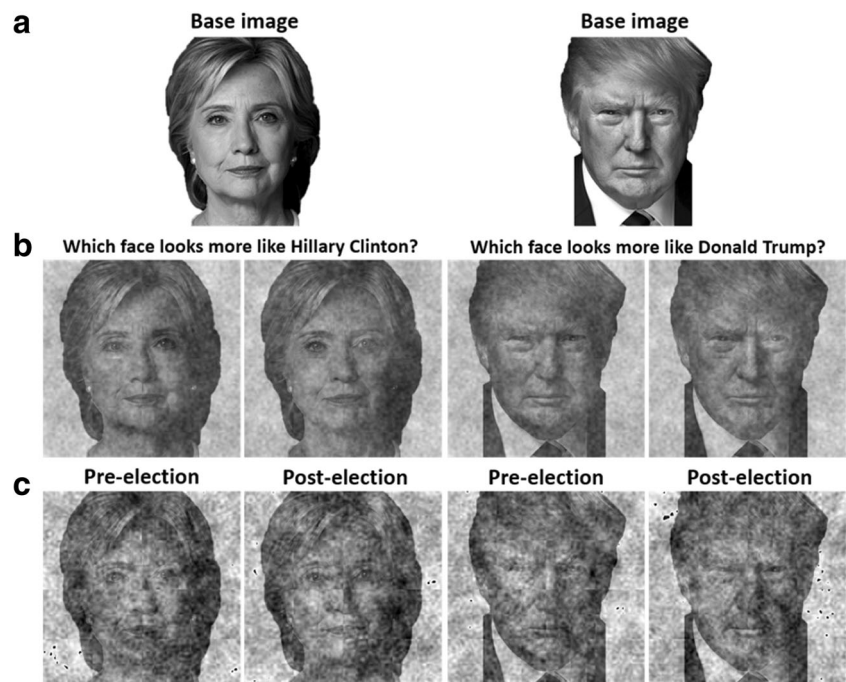
classification phase is that participants may not be aware of the criteria they adopt to endorse a face as looking more like, for example, Hillary Clinton (Brinkman et al. 2017). Here, participants could spontaneously use the information presented to them to select resemblance to Clinton or Trump. A blank screen appeared for 250 ms between trials.

Face Classification Image Processing Pre- and post-election face classification images of Clinton and Trump were generated for each participant by averaging the 100 noise patterns selected by each participant (one per trial) and superimposing that average over the original base image of, respectively, Clinton or Trump (Fig. 1c). This process yielded 228 face classification images (57 each of pre- and post-election Clinton and Trump). Face classification images reflected participants' unique mental representations of candidate faces (e.g., how a participant visualized Clinton's face pre-election).

Face Ratings Phase

Twenty-five individuals from Amazon Mechanical Turk were compensated \$.50 for using a 7-point scale to rate the competence of each classification image in two randomly presented blocks: "How competent does this face look?," rated from 1 (*not at all competent*) to 7 (*extremely competent*). The number of raters was selected on the basis of past and recent reverse correlation work (e.g., Krendl and Freeman 2017). Raters were naïve as to how the faces were generated. Blocks consisted of 114 Clinton or 114 Trump classification images. Ratings of the

Fig. 1 Base images of Hillary Clinton and Donald Trump (a), example reverse correlation task stimuli (b), and pre- and post-election classification images from the same participant (c)



Clinton images ranged from 3.28 to 5.16 ($M = 4.13$, $SD = .36$), and ratings of the Trump images ranged from 2.60 to 4.88 ($M = 3.99$, $SD = .43$). Rated face classification images likely reflect visual reflections of participants' internal representations guiding how they perceive faces (Brinkman et al. 2017). The average competence rating of each classification image thus reflects internal representations of candidate competence unique to participants.

Verifying Competence-Specific Effects

Given our hypothesis that people with higher benevolent sexism would yield face classification images of Clinton that appeared more competent, it was necessary to rule out the possibility that these participants might simply produce clearer images that are more likely to be positively rated on any trait (e.g., trustworthiness). To address this possibility, 20 naïve raters from Amazon Mechanical Turk who did not rate the face classification images on competence were compensated \$.50 to rate the face classification images on trustworthiness using a 7-point scale: "How trustworthy does this face look?," rated from 1 (*not at all trustworthy*) to 7 (*extremely trustworthy*).

Results

Hypothesis 1: Benevolent Sexism and Attitudes toward the 2016 Election Outcome

Our first goal was to illustrate how benevolent sexism ($M = 2.16$, $SD = .94$), hostile sexism ($M = 1.82$, $SD = .96$), or both predicted the maintenance of traditional attitudes via attitudes toward the 2016 election. To this end, we regressed attitudes toward the 2016 election outcome ($M = 3.55$, $SD = 2.50$; higher values reflect more positive attitudes), opposition to Trump's candidacy pre-election ($M = 7.51$, $SD = 2.75$), and opposition to Trump's candidacy post-election ($M = 7.47$, $SD = 2.49$) on benevolent and hostile sexism. Because benevolent sexism relates to maintaining traditional gender roles (Glick and Fiske 2001), we expected benevolent sexism to predict more positive attitudes toward the election outcome and less opposition to Trump's candidacy (Hypothesis 1). We tested if hostile sexism relates to more opposition to Clinton's candidacy pre-election ($M = 6.89$, $SD = 2.61$) and post-election ($M = 6.14$, $SD = 2.73$) because hostile sexism is characterized by the explicit derogation of women (Glick and Fiske 1996, 2001) and thus potentially their candidacies. See Table 1 for regression statistics.

The model predicting attitudes toward the 2016 election outcome from benevolent and hostile sexism was significant, $F(2, 54) = 4.33$, $p = .02$, $f^2 = .16$. Supporting Hypothesis 1, benevolent, but not hostile, sexism predicted more positive

Table 1 Predicting attitudes toward the 2016 election and opposition to Clinton's and Trump's candidacies pre- and post-election from benevolent and hostile sexism, study 1

Model	<i>B</i> (<i>SE</i>)	<i>t</i>	<i>p</i>	<i>R</i>	<i>R</i> ²
Attitudes toward 2016 Election				.37	.14
Benevolent Sexism	.86 (.41)	2.10	.04		
Hostile Sexism	.20 (.40)	.51	.62		
Pre-Election Opposition to Trump				.40	.16
Benevolent Sexism	-1.15 (.45)	2.59	.01		
Hostile Sexism	-.04 (.43)	.09	.93		
Post-Election Opposition to Trump				.52	.27
Benevolent Sexism	-1.50 (.38)	3.99	<.001		
Hostile Sexism	.25 (.37)	.69	.49		
Pre-Election Opposition to Clinton				.39	.15
Benevolent Sexism	.31 (.42)	.72	.47		
Hostile Sexism	.87 (.41)	2.11	.04		
Post-Election Opposition to Clinton				.42	.18
Benevolent Sexism	.67 (.44)	1.53	.13		
Hostile Sexism	.69 (.43)	1.62	.11		

attitudes toward the 2016 election outcome (see Table 1). Also supporting Hypothesis 1 were models predicting opposition to Trump's candidacy pre-election, $F(2, 54) = 5.12$, $p = .01$, $f^2 = .19$, and post-election, $F(2, 54) = 9.74$, $p < .001$, $f^2 = .36$. At both timepoints, benevolent, but not hostile, sexism predicted less opposition to Trump's candidacy.

An additional model predicting opposition to Clinton's candidacy pre-election was significant, $F(2, 54) = 4.92$, $p = .01$, $f^2 = .18$. Here, hostile, but not benevolent, sexism predicted more opposition to Clinton's candidacy (see Table 1). Speculatively, this could be because openly expressing opposition to a woman's candidacy might be more in line with the derogation of women that is consistent with hostile sexism (Glick and Fiske 1996). Although a model predicting opposition to Clinton's candidacy post-election was significant, $F(2, 54) = 5.71$, $p = .01$, $f^2 = .21$, neither benevolent nor hostile sexism predicted it. Notably, only benevolent sexism predicted endorsing Trump's candidacy and favorable attitudes toward the 2016 election outcome. Both findings are consistent with the idea of maintaining traditional gender roles. We thus focused on benevolent sexism when examining its potential relationship with perceived competence of Clinton (see the [online supplement](#) for analyses on hostile sexism).

Hypotheses 2 and 3: Benevolent Sexism and women's Competence

Two key goals of Study 1 were to test for a potentially positive relationship between benevolent sexism and perceived

competence of Clinton (Hypothesis 2) and to determine if this expected relationship was stronger pre- versus post-election (Hypothesis 3). The data collected in Study 1 were inherently multilevel because four classification images were nested within each participant. To address Hypotheses 2 and 3, we used hierarchical linear modeling (HLM; Bryk and Raudenbush 1992; Cassidy and Gutchess 2015) to analyze a data structure where pre- and post- election face classification images of Clinton and Trump (level-1) were nested within participants who completed the face classification phase (level-2). HLM is useful in this context because it tests how competence evaluations of level-1 predictors (i.e., face classification images of Clinton or Trump) vary as a function of level-2 characteristics (e.g., benevolent sexism). Support for Hypothesis 2 would emerge if an interaction between Benevolent Sexism (level-2; grand-mean centered) and Candidate (level-1; coded as 0 = Trump and 1 = Clinton) predicted Competence Ratings of the face classification images (level-1 outcome variable). Support for Hypothesis 3 would emerge if the interaction between Benevolent Sexism and Candidate on Competence Ratings varied by Election (level-1; coded as 0 = Pre and 1 = Post).

To control for potential confounds associated with ideological beliefs (e.g., right-wing authoritarianism and social dominance orientation) that might affect perceptions of Clinton (Choma and Hanoch 2017), we included Election Attitudes pre-election (level-2; grand-mean centered) in the model. Favorable attitudes toward the 2016 presidential election outcome corresponded with ideological beliefs including right-wing authoritarianism and social dominance orientation (Choma and Hanoch 2017), thus providing a theoretical justification for including Election Attitudes in our model. Given their relevance to the hypotheses, all variables were simultaneously entered into the HLM. See Table 2a for HLM statistics.

Supporting that benevolent sexism would positively relate to perceiving competence in Hillary Clinton (Hypothesis 2), an interaction emerged between Benevolent Sexism and Candidate on Competence Ratings of participants' face classification images (see Table 2a). Benevolent Sexism predicted more perceived competence for Clinton, $b = .11$, $SE = .05$, $t(54) = 2.41$, $p = .02$, but not for Trump, $b = -.04$, $SE = .07$, $t(54) = .61$, $p = .55$.

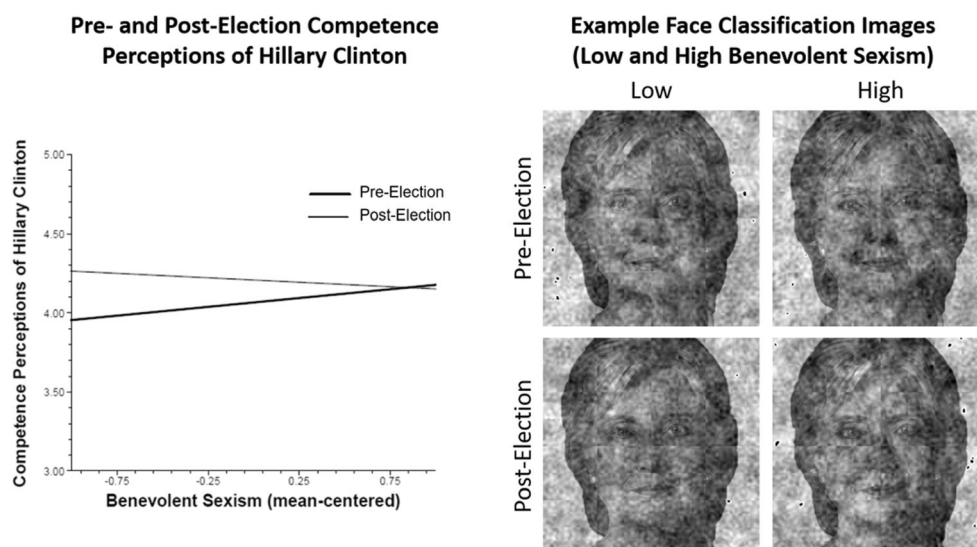
From the lens of shifting standards (Biernat 2003), a positive relationship between benevolent sexism and perceived competence of Clinton may be explained by Clinton's being evaluated against the lower competency standard for women held by people higher in sexism (e.g., Biernat and Manis 1994). As a high achieving woman, Clinton would defy a lower competency stereotype. In a context in which she may be perceived as less competent (e.g., after her loss; ThorOUGHgood et al. 2013), however, Clinton might not defy

Table 2 Predicting competence and trustworthiness perceptions, study 1

Regression models	<i>b</i> (SE)	<i>t</i>	<i>p</i>
(a) Competence			
Intercept	3.93 (.06)	68.79	< .001
Within-Person (Level 1)			
Candidate	.14 (.07)	1.94	.05
Election	.12 (.06)	1.83	.07
Candidate × Election	.03 (.08)	.31	.76
Between-Person (Level 2)			
Benevolent Sexism (BS)	-.04 (.06)	.73	.47
Election Attitudes (EA)	.01 (.02)	.61	.55
Interactions Between Level 1 and Level 2			
Candidate × BS	.15 (.07)	2.04	.04
Candidate × EA	-.03 (.03)	1.17	.24
Election × BS	.07 (.07)	1.02	.31
Election × EA	-.04 (.03)	1.51	.11
Candidate × Election × BS	-.23 (.10)	2.45	.02
Candidate × Election × EA	.08 (.05)	1.49	.14
(b) Trustworthiness			
Intercept	3.39 (.06)	54.72	< .001
Within-Person (Level 1)			
Candidate	.34 (.09)	3.63	<.001
Election	.08 (.08)	1.05	.29
Candidate × Election	-.06 (.13)	.45	.65
Between-Person (Level 2)			
Benevolent Sexism (BS)	-.02 (.08)	.20	.84
Election Attitudes (EA)	.03 (.03)	1.14	.26
Interactions Between Level 1 and Level 2			
Candidate × BS	.07 (.10)	.68	.50
Candidate × EA	-.04 (.04)	.95	.34
Election × BS	.04 (.09)	.45	.66
Election × EA	-.03 (.03)	.95	.34
Candidate × Election × BS	-.06 (.14)	.45	.65
Candidate × Election × EA	.01 (.05)	.25	.80

that stereotype to the same degree and thus not be perceived as especially competent. Supporting Hypothesis 3, the interaction between Benevolent Sexism and Candidate on Competence Ratings of participants' face classification images was qualified by Election (pre- versus post-) ($b = -.23$, $SE = .10$), $t(162) = 2.46$, $p = .02$. Benevolent sexism predicted face classification images of Clinton being rated as more competent pre-election ($b = .11$, $SE = .05$), $t(54) = 2.41$, $p = .02$, but not post-election ($b = -.06$, $SE = .05$), $t(54) = 1.04$, $p = .30$ (see Fig. 2). Benevolent sexism did not predict competence perceived in face classification images of Trump ($b = -.04$, $SE = .05$), $t(54) = .73$, $p = .47$, or post-election ($b = .03$, $SE = .06$), $t(54) = .45$, $p = .66$.

Fig. 2 Regression lines for benevolent sexism effects on competence perceptions of Hillary Clinton pre-election (black line) and post-election (gray line) and example classification images from one participant low in benevolent sexism (.70) and one high in benevolent sexism (3.90) from study 1



Verifying Competence-Specific Effects

We did not find support for the possibility that more perceived competence reflected a broader positivity bias in perceptions of Clinton with more benevolent sexism. Specifically, there were no significant effects of participants' Benevolent Sexism on Trustworthiness Ratings of participants' face classification images (see Table 2b).

Discussion

Study 1 showed that benevolent, and not hostile, sexism predicted more positive attitudes toward Donald Trump's presidential candidacy and the 2016 United States election outcome in which Trump won. One interpretation of this finding is that it conceptually replicates work showing that benevolent, and not hostile, sexism predicts upholding traditional gender roles in masculine positions (Glick and Fiske 2001; Hideg and Ferris 2016; King et al. 2012), such as the presidency (Paul and Smith 2008; Smith et al. 2007). Indeed, favoring men's candidacies maintains the masculine nature of the presidency (Smith et al. 2007) and the stereotype that men are better suited for politics (Bracic et al. 2018) because there has never been a female United States president. These findings notably contrast work showing hostile sexism to predict favorable opinions about Donald Trump (Bock et al. 2017; Ratliff et al. 2017). Past work has largely focused on attitudes toward Trump, the candidate, versus the idea of his candidacy. One possibility is that hostile sexism relates more strongly to attitudes toward a male candidate than the tradition maintained by his candidacy (i.e., that men, not women, should

hold powerful political positions). Although beyond the scope of this work to disentangle, future work may examine this possibility.

At the same time, Study 1 tested for a positive relationship between benevolent sexism and perceptions of Clinton's competence. This relationship did not generalize to trait perceptions of trustworthiness from Clinton's face, meaning benevolent sexism did not elicit broadly more positive trait impressions of her. That benevolent sexism predicted more competent perceptions of Clinton's face is suggestive of theoretical models of shifting standards (Biernat 2003). Past shifting standards work has shown that with more prejudice, people have a larger shift in the baseline level of a trait against which target group members may be compared within their group (Biernat and Manis 1994). Speculatively, that baseline could be a lower competency standard for women among people higher in benevolent sexism. A lower baseline for women's competency would allow benevolent sexists to perceive women as especially competent because she would be compared to other women for whom a stereotype of lower competency would be more applicable. Benevolent sexism did not predict competence perceptions of Trump, speculatively because his characterization might not counter a presumed baseline of higher male competency.

Our finding that benevolent sexism only predicted Clinton's perceived competence pre-election also suggests the possibility that, reflecting shifting standards, her competence was evaluated relative to a baseline of lower female competency. Because Clinton lost the election, she would not defy a stereotype of lower female competency after her loss to the degree that she did when she was expected to win. That is, benevolent sexists might not see her as especially

competent compared to other women given her loss. Such a pattern is consistent with work showing that leaders who commit errors (e.g., mismanage their business) are perceived as being less competent (Thoroughgood et al. 2013). This finding also suggests that representations of traits in faces are not static. Although it is well-known that attitudes affect how traits are perceived in faces (e.g., Dotsch et al. 2008), our finding complements the impression updating literature (e.g., Mende-Siedlecki et al. 2013) by being the first known to suggest that representations of traits in faces may differ based on context.

In Study 1, competence was inferred through independent ratings of face classification images and not participants' explicit evaluations. Using reverse correlation to test for a relationship between benevolent sexism and perceived competence of Clinton was beneficial because it allowed for evaluations of Clinton to be potentially less affected by desirability to be consistent with partisan beliefs (as in Wright and Tomlinson 2018). However, a limitation of reverse correlation is that it cannot elucidate the group against which Clinton's competence was evaluated (e.g., if she was compared to other women). Further, although reverse correlation is widely used to estimate trait perceptions in faces (e.g., Dotsch et al. 2008; Ratner et al. 2014; Young et al. 2014), we cannot rule out that representations of Clinton's face reflected participants' own perceptions of her competency versus media depictions that were more goal-relevant for certain perceivers. Finally, Hillary Clinton and Donald Trump are so well known and polarizing that it is possible that they are not truly representative of people in prominent positions. These limitations make it unclear if the positive relationship between benevolent sexism and perceived competence shown in Study 1 extends to other women in prominent positions. To address these limitations, we manipulated the gender against which an unidentified senator's competence would be evaluated and obtained explicit competence evaluations of this senator in Study 2.

Study 2

Study 2 examined if people higher in benevolent sexism use shifting standards when they evaluate a woman in a prominent leadership position. Specifically, the goal of Study 2 was to determine if benevolent sexism positively related to perceiving women as more competent when evaluating them against other women versus against other men. The extent to which women are perceived as more competent against women than they are against men would reflect the extent of shifting standards (Biernat and Manis 1994). We tested for this possibility using explicit competence evaluations of unidentified female and male U.S. senators. Critically, we manipulated the

reference category (i.e., evaluating senators against men or women) used when making competence evaluations. If people higher in benevolent sexism have a lower competency standard for women, benevolent sexism should positively relate to the extent of shifting standards when evaluating a woman's competency. That is, benevolent sexism should positively relate to evaluating a woman as more competent relative to other women versus other men (Hypothesis 1).

At the same time, we tested if people higher in benevolent sexism nevertheless maintained traditional gender roles. In Study 1, this idea was reflected in benevolent sexism predicting favorable attitudes toward the 2016 presidential election outcome in which a man was successful. In Study 2, we extended this idea by examining if benevolent sexism positively related to men or negatively related to women being expected to be extremely successful as a senator (Hypothesis 2). Finally, we also tested if the extent of shifting standards when evaluating women mediated a relationship between benevolent sexism and expectations of women's success to link shifting standards to both benevolent sexism and negative outcomes for women. These patterns would extend the literature by showing that stronger benevolent sexists acknowledge women's high achievements by evaluating them as especially competent (via shifting standards), yet ultimately have expectations of men, and not of women, being successful in prominent positions.

Method

Participants

Two hundred individuals recruited from Amazon Mechanical Turk participated. This sample size was chosen to ensure that usable data emerged from at least 30 participants who were either high or low in benevolent sexism and who evaluated a male or female senator (e.g., at least 120 participants total; see Results). Participants provided informed consent and were compensated \$.50. Eleven adults were excluded for failing a manipulation check, yielding 189 analyzed adults ($M_{age} = 38.42$ years, $SD = 12.25$, 86 female).

Procedure

Participants were randomly assigned to evaluate a female senator ($n = 90$) or a male senator ($n = 99$). A Chi-square test showed male and female participants were evenly distributed across conditions, $\chi^2(1, n = 189) = 1.34, p = .25$. We manipulated Target Gender between-subjects because we did not want evaluations of one target to influence evaluations of a new target, potentially clouding Target Gender interactions with other variables. Participants were told they would be

evaluating the senator's job performance. We next oriented participants to thinking about job performance relative to others using a method from shifting standards work (e.g., Biernat and Manis 1994). Participants were told:

Think about people who are United States senators. Now think about 100 women [men] in the population. Some of these people will do well in this profession, and some will not. Please distribute these 100 women [men] into these bins based on the probability of their doing well in this profession.

The levels of the bins were “extremely unlikely,” “moderately likely,” “slightly unlikely,” “neither likely nor unlikely,” “slightly likely,” “moderately likely,” and “extremely likely.” That is, participants might mark 100 in the “extremely unlikely” bin or mark a more even distribution of women across the bins. Although there were seven bins, these bins did not constitute a continuous variable. Our hypotheses regarded the effects of benevolent sexism on expectations of men and women being extremely likely to be successful in prominent political positions. Analyses thus focused on the bin best reflecting that idea. Analyses of other bins were not conducted because they were either irrelevant to (e.g., being binned as slightly likely to be successful) or redundant with (e.g., being binned as extremely unlikely to be successful) our research questions. It was important for participants to distribute among bins to be consistent with related work similarly orienting participants to thinking about performance relative to others (Biernat and Manis 1994). Participants were then told, “For the next several questions, I want you to imagine a female [male] senator who is likely to be successful in that position.”

To manipulate a reference category within-subjects, and thus to determine if benevolent sexism related to higher competence evaluations of women relative to other women versus men (i.e., shifting standards), the next four randomly presented questions involved evaluating the target's competence relative to other women or other men using 7-point scales. Manipulating reference category within-subjects was important because it allowed us to determine how the same participants evaluated a senator when placed in different contexts (i.e., evaluated against women versus men). Two questions evaluated the target's competence relative to other women (e.g., “Relative to other women, how competent would you expect this person to be when performing the responsibilities of a senator?” where 1 = not at all competent to 7 = very competent; “Relative to other women, at what percentage of the responsibilities of a senator would you expect this person to be more competent?” where 1 = < 10% to 7 = > 90%) and two questions evaluated the target's competence relative to other men. The order of these questions was randomized across participants. The two responses evaluating competence

relative to other women were correlated and reliable, $r(187) = .65, p < .001$ (Cronbach's $\alpha = .78$), and the two responses evaluating competence relative to other men were correlated and reliable, $r(187) = .55, p < .001$ (Cronbach's $\alpha = .70$). We thus averaged the two evaluations relative to other women and the two evaluations relative to other men for all analyses. Competence evaluations of the female target relative to women versus relative to men reflected the extent of shifting standards and served as the dependent variable in the following analyses.

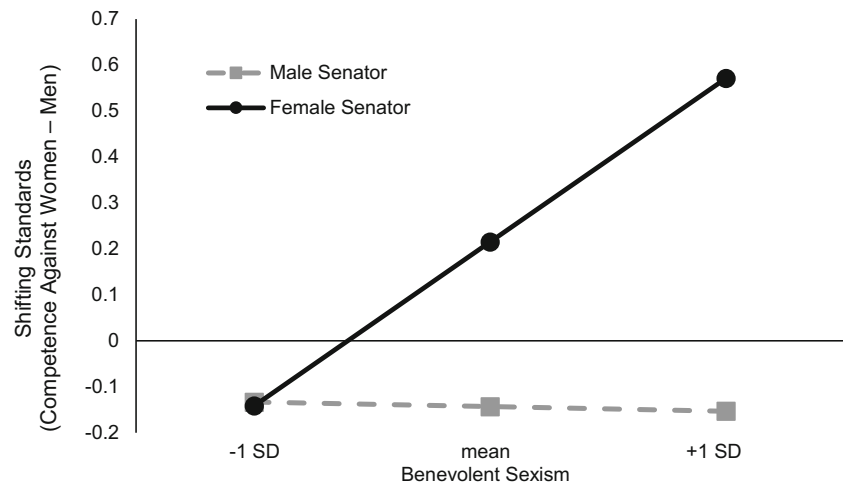
Participants then indicated the percentage of senators they believed were male using a scale ranging from 0% to 100% in 10% increments and completed the ASI. Responses on the ASI were reliable (Hostile sexism Cronbach's $\alpha = .92$; Benevolent sexism Cronbach's $\alpha = .88$). Like Study 1, participants reported more benevolent ($M = 2.24, SD = 1.10$) than hostile ($M = 1.93, SD = 1.17$) sexism, $t(188) = 3.89, p < .001, d = .28$. Benevolent and hostile sexism were also correlated, $r(187) = .52, p < .001$. Male participants ($M = 2.20, SD = 1.11$) reported more hostile sexism than did female participants ($M = 1.61, SD = 1.17$), $t(187) = 3.57, p < .001, d = .52$. Male ($M = 2.37, SD = 1.02$) and female ($M = 2.09, SD = 1.17$) participants did not differ in their benevolent sexism, $t(187) = 1.79, p = .08, d = .26$. Because these data suggest some gender differences in sexism, we modeled participants' gender in the following analyses to test for gender effects beyond our a priori hypotheses. Participants believed that 76.83% ($SD = 11.78$) of senators are male, consistent with political offices being predominantly held by men. When data were collected (October 2017), 79% of U.S. senators were male. Lastly, participants identified the target individual (e.g., female senator) to ensure they had paid attention to the task. As we noted previously, 11 participants who responded incorrectly were excluded.

Results

Hypothesis 1: Benevolent Sexism Will Positively Relate to Shifting Standards toward Women

Shifting standards is defined as the extent to which a target is evaluated differently against one group versus another. We therefore created a difference score to isolate the extent to which participants evaluated a woman as competent relative to other women versus the extent to which they evaluated her as competent relative to other men. To examine if benevolent sexism related to more shifting standards in evaluations of women's competence, we regressed competence evaluations of women versus men on Target Gender (coded as 0 = man and 1 = woman), Benevolent Sexism (mean centered), and their interaction. The model was significant, $F(3, 185) = 5.19, p = .002, f^2 = .08, R^2 = .08$. There was no effect of benevolent sexism ($b = -.009, SE = .10, t = .09, p = .93$). There

Fig. 3 In study 2, benevolent sexism predicted the extent of shifting standards in competence evaluations of women, and not men, in a prominent position



was an effect of Target Gender ($b = .36$, $SE = .16$, $t = 2.18$, $p = .03$), suggesting more shifting standards when evaluating female versus male targets.

Supporting Hypothesis 1, an interaction between Benevolent Sexism and Target Gender qualified these effects ($b = .33$, $SE = .15$, $t = 2.22$, $p = .03$) (see Fig. 3). When evaluating female target senators, benevolent sexism positively related to evaluating targets as more competent against women versus men ($b = .32$, $SE = .11$, $t = 2.98$, $p = .003$). That is, participants higher in benevolent sexism evaluated the same woman as being more competent when she was compared to other women than when she was compared to men. When evaluating male targets, benevolent sexism did not relate to competence evaluations against women versus men ($b = -.009$, $SE = .10$, $t = .08$, $p = .93$). (See the [online supplement](#) for regressions on evaluations against other women and against other men, as well as for analyses using hostile sexism.)

Unlike Study 1, male and female participants showed some differences in sexism. To rule out that the present effects differed by participant gender, we included Participant Gender and its interactions with the other variables in a second model. The second model did not account for more variance than the first model (R^2 change = .01).

Hypothesis 2: Benevolent Sexism Will Relate to Expectations of Male Success

In Study 1, benevolent sexism predicted favorable attitudes toward an outcome in which a man was successful and the masculine nature of the presidency was maintained (e.g., Paul and Smith 2008; Smith et al. 2007). To complement this finding, we examined if benevolent sexism related to expectations of men's and women's success as senators. Here, we examined the number of men and women expected to do extremely well as senators. Supporting Hypothesis 2, benevolent sexism corresponded with more men binned as extremely likely to do

well, $r(97) = .24$, $p = .02$. Benevolent sexism, however, was not significantly related to women being binned as being extremely likely to do well, $r(88) = .10$, $p = .34$.

It may seem counterintuitive that benevolent sexism was not significantly related to expectations of women being successful as senators. Shifting standards, however, might allow for people higher in benevolent sexism to evaluate women as highly competent yet not expect their success at the same time. To address this possibility, we tested if the extent to which people evaluated women as more being competent relative to women versus men (i.e., the extent of shifting standards) mediated a relationship between benevolent sexism and women being binned as extremely likely to do well as a senator (see Fig. 4). A non-significant total effect does not prohibit testing for an indirect effect (Hayes 2009). We conducted this analysis using PROCESS for SPSS (Hayes 2012) with 5000 bootstrap samples for bias-corrected confidence intervals.

Shifting standards mediated a relationship between lower expectations of women's success and evaluations of competence ($b = -1.37$, $SE = .79$, 95% CI [-3.24, -.21]). Specifically, even though benevolent sexism did not predict the likelihood that women were binned as being extremely

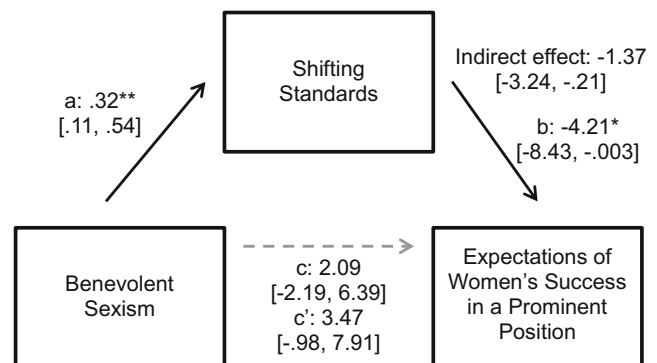


Fig. 4 Shifting standards mediated a relationship between benevolent sexism and expectations of women's success as a senator. Coefficients are unstandardized. Values in brackets are 95% confidence intervals. * $p < .05$. ** $p < .01$

likely to do well as a senator ($b = 2.09$, $SE = 2.16$, $t = .97$, $p = .33$, 95% CI [-2.19, 6.39]), it positively related to the extent to which shifting standards were used to evaluate her competence ($b = .32$, $SE = .11$, $t = 3.04$, $p = .003$, 95% CI [.11, .54]). Benevolent sexism also predicted the extent to which shifting standards negatively related to women being binned as being extremely likely to do well as a senator ($b = -4.21$, $SE = 2.12$, $t = 1.99$, $p = .049$, 95% CI [-8.43, -.003]). Benevolent sexism thus related to lower expectations of women's success through more use of shifting standards to evaluate her competence.

Discussion

Study 1 tested for and found a positive relationship between benevolent sexism and competence evaluations of prominent women. Study 2 conceptually replicated and elaborated on this finding by showing that benevolent sexism positively related to evaluating a female senator as more competent when she was evaluated relative to other women than relative to other men.

Study 2 suggests shifting standards as a process underlying the key finding of Study 1. Because the category against which a target was evaluated was explicitly manipulated, we could show that benevolent sexism positively related to a female senator being evaluated as more competent when she was evaluated against a group stereotyped to be less competent. This finding is important because it may illustrate one way benevolent sexists are unlikely to be perceived as prejudiced (Barreto and Ellemers 2005): Benevolent sexists may praise a high achieving woman as being especially competent given a lower baseline against which to compare her. This finding also informs Study 1 by supporting the possibility that Clinton was evaluated as especially competent by people higher in benevolent sexism because she was evaluated against other women. Higher benevolent sexism did not yield male senators being evaluated as more competent against women versus against men. Speculatively, this could be because benevolent sexism is related to chivalry more so than antagonism toward women (Glick and Fiske 1997). That is, rating a man as more competent relative to women than to men could be construed as antagonistic. Alternatively, participants may feel as though explicitly endorsing a man as especially competent against women as compared to against men is not socially desirable and avoid that response. Future work may disentangle these possibilities.

Although showing that benevolent sexism positively related to using shifting standards to evaluate women's competence, Study 2 also showed benevolent sexism related to expecting more male success and less female success. Here, higher benevolent sexism positively related to men being expected to do extremely well as a senator. Despite higher

benevolent sexists' evaluations of female senators as especially competent in some circumstances (i.e., when evaluated against women), higher benevolent sexism is nevertheless associated with expectations of male success. Further, whereas benevolent sexists used shifting standards more to evaluate women's competence, greater use of shifting standards related to expecting fewer women to be successful as senators. The present data provide initial evidence that evaluating women as especially competent relative to other women (instead of relative to men) may be a way for people higher in benevolent sexism to outwardly praise some women yet ultimately expect them to be less successful.

General Discussion

Over two studies, we examined how benevolent sexism affects competence evaluations of women in prominent political positions. Women are stereotyped to be less competent than men are (Broverman et al. 1972). Reflecting shifting standards, stereotypes lower the standard more prejudiced people use to evaluate stereotyped group members (Biernat and Manis 1994). Evaluating a woman's competence using shifting standards may allow a woman to be perceived as especially competent (relative to a low baseline) by people who more strongly endorse benevolent sexism. Using different methods, Studies 1 and 2 supported this possibility. This possibility, however, would not preclude higher benevolent sexists from supporting men for prominent positions or expecting them to be more successful. Indeed, favorable attitudes and expectations toward men reaching prominent positions emerged across both studies. Linking shifting standards to expectations of women's success, shifting standards mediated a relationship between benevolent sexism and expectations of women's success in Study 2. This finding suggests that shifting standards may allow stronger benevolent sexists to outwardly praise women yet maintain stereotypic expectations of their lower success.

Study 1 showed that benevolent sexism predicted lower opposition to Donald Trump's presidential candidacy and more favorable attitudes toward the 2016 United States election outcome. Paralleling Study 1, Study 2 showed benevolent sexism to be positively associated with men being expected to be successful as senators. These patterns conceptually replicated work showing that benevolent sexism relates to maintaining gender roles (e.g., Hideg and Ferris 2016) through upholding a tradition of men in prominent political offices (Paul and Smith 2008; Smith et al. 2007). At the same time, Study 1 found a positive relationship between benevolent sexism and the perceived competence of Hillary Clinton. Suggesting that this pattern reflected shifting standards, Study 2 showed that benevolent sexism positively related to evaluating a female senator as especially competent when she

was evaluated against other women relative to when she was evaluated against other men. Study 2 supports the possibility that benevolent sexism positively related to perceptions of Clinton's competence because she was naturally evaluated against a within-group standard of other women versus a standard encompassing several groups or men alone.

Study 2 also showed that the extent to which women are evaluated as being more competent when compared to other women than when compared to other men related to both benevolent sexism and expectations of women's success. Using shifting standards when evaluating prominent women may allow benevolent sexists to praise these women as competent and maintain lower expectations of their success at the same time. Speculatively, lower expectations may allow for traditional gender roles to be maintained through the endorsement of men's candidacies. Indeed, gendered stereotypes that men are better suited for politics that positively relate to supporting male (versus female) candidates (Bracic et al. 2018) could elicit more use of shifting standards when evaluating women for political positions. Future work might directly address this possibility.

Our findings extend the literatures on benevolent sexism and shifting standards. Speaking to the former, our data suggest that benevolent sexists might not have more positive characterizations of women (Glick and Fiske 2001). Instead, benevolent sexists might be more likely to use shifting standards to maintain positive characterizations while upholding traditional gender roles. Speaking to the latter, the present findings conceptually replicate work showing that higher prejudiced individuals use shifting standards more (Biernat and Manis 1994), and they show how shifting standards may reconcile positive attitudes toward stereotyped group members with stereotypic expectations for them. Women in prominent political positions may be evaluated as highly competent because they are compared to a lower within-group competency standard.

Limitations

A key strength of Study 1 was its ecological validity. However, although a positive relationship between benevolent sexism and Hillary Clinton's perceived competence is theoretically consistent with her being evaluated against other women, a limitation of Study 1 was that it did not determine the group against which she was evaluated. Study 2 addressed this limitation using an explicit manipulation of the group against which a female senator would be evaluated. It will be important, however, for future work to develop strategies to determine if shifting standards actually elicit especially competent evaluations of specific prominent women because more women are garnering attention on the national stage. A second limitation regards that women in high political power vary in the degree to which they defy traditional gender roles.

Although Hillary Clinton and Sarah Palin (a U.S. vice presidential nominee in 2008) were both candidates for high political offices, for instance, Clinton was perceived to defy traditional gender roles more so than Palin (Carlin and Winfrey 2009; Gervais and Hillard 2011). Interestingly, benevolent sexism positively related to support for Palin in 2008 even though Clinton was perceived as more competent (Gervais and Hillard 2011). These findings raise the possibility that the mechanism discussed in the current work could be more applicable to women in prominent positions who defy gender roles versus women in prominent positions overall. It will be important for future work to examine the reach of the present findings.

Future Directions

Evaluating prominent women against other women might allow for a discrepancy in how people higher in benevolent sexism reconcile a woman's high achievements with not supporting her. Another factor potentially contributing to this discrepancy is the media's focus on the novelty of having women in leadership positions rather than on their qualifications, a factor undermining their legitimacy (Meeks 2013). This undermining may be particularly important for women striving for male-dominated positions because perceptions of these women as novel may overshadow their competency. Indeed, the male-dominated history of the presidency suggests that a man "should be" president, meaning that female candidates face an uphill battle because their gender contrasts how the position has been traditionally situated (Prentice and Carranza 2002).

Further, the gendered nature of positions like the U.S. presidency may yield backlash toward female candidates. Although prominent women may, in part, overcome the stereotype of being less competent than men are, they may face penalties for behaving in counter-stereotypic ways (Phelan et al. 2008), a pattern prevalent among female leaders (Rudman et al. 2012). Indeed, women who succeed in male-dominated professions are perceived as competent, but also as lacking warmth (Eagly and Karau 2002; Heilman et al. 2004). These findings suggest that sexism may relate to negative outcomes due to perceiving prominent women as lacking warmth rather than because they are evaluated against other women as suggested by the present work. Trait perceptions of trustworthiness (a trait closely related to warmth; Fiske et al. 2007) in Hillary Clinton's face, however, were not predicted by benevolent sexism in Study 1. It will be important for future work to examine backlash effects in their connection to the effects of comparing prominent women to other women on evaluating competence.

Finally, the present work focused on women in traditionally masculine positions illustrating shifting standards as way for people higher in benevolent sexism to evaluate these women

as especially competent. At the same time, benevolent sexism predicted support for men. It is possible that benevolent sexism may only elicit this discrepancy for traditionally masculine positions like the U.S. presidency (Smith et al. 2007). When a position involves leadership but does not defy traditional gender roles (e.g., kindergarten teacher), benevolent sexism may predict especially competent evaluations of women and support for women at the same time. It will be important for future work to examine how the nature of a prominent position may lead to discrepant versus convergent evaluations of competence and support of women.

Practice Implications

The present work may be relevant to vetting political candidates and understanding gender inequality in the workplace. Illustrating that benevolent sexists evaluate prominent women as especially competent when compared to women versus men is important because simply stating that a woman is very competent may lead people to question if sexism still contributes to prominent women not receiving support (e.g., Hillary Clinton's electoral defeats; Lawless 2009). Because hiring committee members may agree that a female candidate is competent, for example, they may not feel as though sexism contributed to a male candidate being chosen for a position. The present work suggests that being evaluated as very competent will elicit less support if evaluations are based on comparisons to people stereotyped to be less competent—evaluations affected by sexism. Indeed, although women may be more likely to make a short list for a stereotypically masculine job, they are less likely to be hired for that job (Biernat and Fuegen 2001). Women may be initially evaluated as especially competent and make short lists because they are more likely to be compared to other women. This tendency still allows for the stereotypic belief that women are less competent overall and could result in less likelihood of hiring, among other reasons (e.g., Eagly and Karau 2002). The present studies point to shifting standards as a potential route by which prominent women may be acknowledged yet ultimately overlooked. Future work may examine how these findings translate into gender inequality in both representation and in the workplace and the political stage.

Conclusion

The current studies further our understanding of how benevolent sexism impacts competence evaluations of women in prominent positions. Our studies suggest the possibility that evaluations of women as especially competent may not be beneficial even though they may seem to be on a superficial level. It will be important for future work to consider these effects of benevolent sexism when developing strategies to reduce the negative effects of sexism on women in everyday life.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

Research Involving Human Participants All studies in this work were approved by the Indiana University Institutional Review Board.

Informed Consent All participants in this work provided informed consent.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

- Ajzen, I., & Fishbein, M. (1977). Attitude-behavior relations: A theoretical analysis and review of empirical research. *Psychological Bulletin*, 84(5), 888–918. <https://doi.org/10.1037//0033-2909.84.5.888>.
- Ajzen, I., & Fishbein, M. (2000). Attitudes and the attitude-behavior relation: Reasoned and automatic processes. *European Review of Social Psychology*, 11(1), 1–33. <https://doi.org/10.1080/14792779943000116>.
- Barreto, M., & Ellemers, N. (2005). The burden of benevolent sexism: How it contributes to the maintenance of gender inequalities. *European Journal of Social Psychology*, 35, 633–642. <https://doi.org/10.1002/ejsp.270>.
- Becker, J., & Wright, S. (2011). Yet another dark side of chivalry: Benevolent sexism undermines and hostile sexism motivates collective action for social change. *Journal of Personality and Social Psychology*, 101(1), 62–77. <https://doi.org/10.1037/a0022615>.
- Biernat, M. (2003). Toward a broader view of social stereotyping. *American Psychologist*, 58(12), 1019–1027. <https://doi.org/10.1037/0003-066x.58.12.1019>.
- Biernat, M., & Fuegen, K. (2001). Shifting standards and the evaluation of competence: Complexity in gender-based judgment and decision-making. *Journal of Social Issues*, 57, 707–724. <https://doi.org/10.1111/0022-4537.00237>.
- Biernat, M., & Kobrynowicz, D. (1997). Gender- and race-based standards of competence: Lower minimum standards but higher ability standards for devalued groups. *Journal of Personality and Social Psychology*, 72(3), 544–557. <https://doi.org/10.1037/0022-3514.72.3.544>.
- Biernat, M., & Manis, M. (1994). Shifting standards and stereotype-based judgments. *Journal of Personality and Social Psychology*, 66(1), 5–20. <https://doi.org/10.1037//0022-3514.66.1.5>.
- Biernat, M., Manis, M., & Nelson, T. (1991). Stereotypes and standards of judgment. *Journal of Personality and Social Psychology*, 60(4), 485–499. <https://doi.org/10.1037//0022-3514.60.4.485>.
- Bock, J., Byrd-Craven, J., & Burkley, M. (2017). The role of sexism in voting in the 2016 presidential election. *Personality and Individual Differences*, 119, 189–193. <https://doi.org/10.1016/j.paid.2017.07.026>.
- Bracic, A., Israel-Trummel, M., & Shortle, A. (2018). Is sexism for white people? Gender stereotypes, race, and the 2016 presidential election. *Political Behavior*, 1–27. <https://doi.org/10.1007/s11109-018-9446-8>.
- Brinkman, L., Todorov, A., & Dotsch, R. (2017). Visualizing mental representations: A primer on noise-based reverse correlation in social psychology. *European Review of Social Psychology*, 28(1), 333–361. <https://doi.org/10.1080/10463283.2017.1381469>.

- Broverman, I., Vogel, R., Broverman, D., Clarkson, T., & Rosenkrantz, P. (1972). Sex-role stereotypes: A current appraisal. *Journal of Social Issues, 28*, 59–78. <https://doi.org/10.1111/j.1540-4560.1972.tb00018.x>.
- Bryk, A., & Raudenbush, S. (1992). *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks: Sage Publications, Inc.
- Carlin, D., & Winfrey, K. (2009). Have you come a long way, baby? Hillary Clinton, Sarah Palin, and sexism in the 2008 campaign coverage. *Communication Studies, 60*, 326–343. <https://doi.org/10.1080/10510970903109904>.
- Cassidy, B., & Gutchess, A. (2015). Neural responses to appearance-behavior congruity. *Social Cognition, 43*(3), 211–226. <https://doi.org/10.1521/soco.2015.33.3.1>.
- Catalyst. (2013). *2013 Catalyst census: Financial post 500 women board directors*. New York: Author.
- Choma, B., & Hanoch, Y. (2017). Cognitive ability and authoritarianism: Understanding support for trump and Clinton. *Personality and Individual Differences, 106*, 287–291. <https://doi.org/10.1016/j.paid.2016.10.054>.
- Darweesh, A., & Abdullah, N. (2016). A critical discourse analysis of Donald Trump's sexist ideology. *Journal of Education and Practice, 7*(30), 87–95. <https://doi.org/10.7575/aic.ijalel.v.6n.5p.1>.
- Dotsch, R., & Todorov, A. (2012). Reverse correlating social face perception. *Social Psychological and Personality Science, 3*(5), 562–571. <https://doi.org/10.1177/1948550611430272>.
- Dotsch, R., Wigboldus, D., Langner, O., & van Knippenberg, A. (2008). Ethnic out-group faces are biased in the prejudiced mind. *Psychological Science, 19*(10), 978–980. <https://doi.org/10.1111/j.1467-9280.2008.02186.x>.
- Dovidio, J., & Gaertner, S. (1998). On the nature of contemporary prejudice: The causes, consequences, and challenges of aversive racism. In S. Fiske & J. Eberhardt (Eds.), *Racism: The problem and the response* (pp. 3–32). Newbury Park: Sage.
- Eagly, A., & Karau, S. (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>.
- Faul, F., Erdfelder, E., Lang, A., & 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.
- Fiske, S., Cuddy, A., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences, 11*(2), 77–83. <https://doi.org/10.1016/j.tics.2006.11.005>.
- Gaffney, A., & Blaylock, D. (2010). Hillary Clinton's race: Did she match the presidential prototype. *Advancing Women in Leadership, 30*(6), 1–15. <https://doi.org/10.18738/awl.v30i0.294>.
- Gervais, S., & Hillard, A. (2011). A role congruity perspective on prejudice toward Hillary Clinton and Sarah Palin. *Analyses of Social Issues and Public Policy, 11*(1), 221–240. <https://doi.org/10.1111/j.1530-2415.2011.01263.x>.
- Glick, P., & Fiske, S. (1996). The ambivalent sexism inventory: Differentiating hostile and benevolent sexism. *Journal of Personality and Social Psychology, 70*(3), 491–512. <https://doi.org/10.1037/0022-3514.70.3.491>.
- Glick, P., & Fiske, S. (1997). Hostile and benevolent sexism: Measuring ambivalent sexist attitudes toward women. *Psychology of Women Quarterly, 1997*(21), 119–135. <https://doi.org/10.1111/j.1471-6402.1997.tb00104.x>.
- Glick, P., & Fiske, S. (2001). An ambivalent alliance: Hostile and benevolent sexism as complementary justifications for gender inequality. *American Psychologist, 56*(2), 109–118. <https://doi.org/10.1017/cbo9781139022736.005>.
- Glick, P., Fiske, S., Mladinic, A., Saiz, J., Abrams, D., Masser, B., ... Lopez, W. (2000). Beyond prejudice as simple antipathy: Hostile and benevolent sexism across cultures. *Journal of Personality and Social Psychology, 79*(5), 763–775. <https://doi.org/10.1037/0022-3514.79.5.763>.
- Hayes, A. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication Monographs, 76*(4), 408–420. <https://doi.org/10.1080/03637750903310360>.
- Hayes, A. (2012). *PROCESS: A versatile computational tool for observed variable mediation, moderation, and conditional process modeling*. Retrieved from <http://www.afhayes.com/public/process2012.pdf>.
- Helman, E., Flake, J., & Freeman, J. (2015). Static and dynamic facial cues differentially affect the consistency of social evaluations. *Personality and Social Psychology Bulletin, 41*(8), 1123–1134. <https://doi.org/10.1177/0146167215591495>.
- Heilman, M., Wallen, A., Fuchs, D., & Tamkins, M. (2004). Penalties for success: Reactions to women who succeed at male gender-typed tasks. *Journal of Applied Psychology, 89*(3), 416–427. <https://doi.org/10.1037/0021-9010.89.3.416>.
- Hideg, I., & Ferris, D. (2016). The compassionate sexist? How benevolent sexism promotes and undermines gender equality in the workplace. *Journal of Personality and Social Psychology, 111*(5), 706–727. <https://doi.org/10.1037/pspi0000072>.
- Huddy, L., & Terkildsen, N. (1993). Gender stereotypes and the perception of male and female candidates. *American Journal of Political Science, 37*(1), 119–147. <https://doi.org/10.2307/2111526>.
- Husnu, S. (2016). The role of ambivalent sexism and religiosity in predicting attitudes toward childlessness in Muslim undergraduate students. *Sex Roles, 75*(11–12), 573–582.
- Jost, J., & Kay, A. (2005). Exposure to benevolent sexism and complementary gender stereotypes. *Journal of Personality and Social Psychology, 88*(3), 498–509. <https://doi.org/10.2139/ssrn.386981>.
- King, E., Botsford, W., Hebl, M., Kasama, S., Dawson, J., & Perkins, A. (2012). Benevolent sexism at work: Gender differences in the distribution of challenging developmental experiences. *Journal of Management, 38*(6), 1835–1866. <https://doi.org/10.1177/0149206310365902>.
- Krendl, A., & Freeman, J. (2017). Are mental illnesses stigmatized for the same reasons? Identifying the stigma-related beliefs underlying common mental illnesses. *Journal of Mental Health, 1–9*. <https://doi.org/10.1080/09638237.2017.1385734>.
- Lawless, J. (2009). Sexism and gender bias in election 2008: A more complex path for women in politics. *Politics & Gender, 5*(1), 70–80. <https://doi.org/10.1017/s1743923x09000051>.
- Lepore, L., & Brown, R. (1997). Category and stereotype activation: Is prejudice inevitable? *Journal of Personality and Social Psychology, 72*(2), 275–287. <https://doi.org/10.1037/0022-3514.72.2.275>.
- Mangini, M., & Biederman, I. (2004). Making the ineffable explicit: Estimating the information employed for face classifications. *Cognitive Science, 28*, 209–226. https://doi.org/10.1207/s15516709cog2802_4.
- Meeks, L. (2013). All the gender that's fit to print: How the New York times covered Hillary Clinton and Sarah Palin in 2008. *Journalism and Mass Communication Quarterly, 90*(3), 520–539. <https://doi.org/10.1177/1077699013493791>.
- Mende-Siedlecki, P., Baron, S., & Todorov, A. (2013). Diagnostic value underlies asymmetric updating of impressions in the morality and ability domains. *The Journal of Neuroscience, 33*(50), 19406–19415. <https://doi.org/10.1523/jneurosci.2334-13.2013>.
- Paul, D., & Smith, J. (2008). Subtle sexism? Examining vote preferences when women run against men for the presidency. *Journal of Women, Politics, and Policy, 29*(4), 451–476. <https://doi.org/10.1080/15544770802092576>.
- Phelan, J., Moss-Racusin, C., & Rudman, L. (2008). Competent yet out in the cold: Shifting criteria for hiring reflect backlash toward agentic women. *Psychology of Women Quarterly, 32*, 406–413. <https://doi.org/10.1111/j.1471-6402.2008.00454.x>.
- Prentice, D., & Carranza, E. (2002). What women should be, shouldn't be, are allowed to be, and don't have to be: The contents of

- prescriptive gender stereotypes. *Psychology of Women Quarterly*, 26, 269–281. <https://doi.org/10.1111/1471-6402.t01-1-00066>.
- Ratliff, K., Redford, L., Conway, J., & Smith, C. (2017). Engendering support: Hostile sexism predicts voting for Donald Trump over Hillary Clinton in the 2016 U.S. presidential election. *Group Processes & Intergroup Relations*. Advance online publication. <https://doi.org/10.1177/1368430217741203>.
- Ratner, K., Dotsch, R., Wigboldus, D., van Knippenberg, A., & Amodio, D. (2014). Visualizing minimal ingroup and outgroup faces: Implications for impressions, attitudes, and behavior. *Journal of Personality and Social Psychology*, 106(6), 897–911. <https://doi.org/10.1037/a0036498>.
- Rosenwasser, S., & Dean, N. (1989). Gender role and political office: Effects of perceived masculinity/femininity of candidate and political office. *Psychology of Women Quarterly*, 13(1), 77–85. <https://doi.org/10.1111/j.1471-6402.1989.tb00986.x>.
- Rudman, L., Moss-Racusin, C., Phelan, J., & Nauts, S. (2012). Status incongruity and backlash effects: Defending the gender hierarchy motivates prejudice against female leaders. *Journal of Experimental Social Psychology*, 48, 165–179. <https://doi.org/10.1016/j.jesp.2011.10.008>.
- Smith, J., Paul, D., & Paul, R. (2007). No place for a woman: Evidence for gender bias in evaluations of presidential candidates. *Basic & Applied Social Psychology*, 29(3), 225–233. <https://doi.org/10.1080/01973530701503069>.
- Thoroughgood, C., Sawyer, K., & Hunter, S. (2013). Real men don't make mistakes: Investigating the effects of leader gender, error type, and the occupational context on leader error perceptions. *Journal of Business and Psychology*, 28(1), 31–48. <https://doi.org/10.1007/s10869-012-9263-8>.
- Todorov, A., Mandisodza, A., Goren, A., & Hall, C. (2005). Inferences of competence from faces predict election outcomes. *Science*, 308(5728), 1623–1626. <https://doi.org/10.1126/science.1110589>.
- Visser, B., Book, A., & Volk, A. (2017). Is Hillary dishonest and Donald narcissistic? A HEXACO analysis of the presidential candidates' public personas. *Personality and Individual Differences*, 106, 281–286. <https://doi.org/10.1016/j.paid.2016.10.053>.
- Wittenbrink, B., Judd, C., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures. *Journal of Personality and Social Psychology*, 72(2), 262–274. <https://doi.org/10.1037//0022-3514.72.2.262>.
- Wright, J., & Tomlinson, M. (2018). Personality profiles of Hillary Clinton and Donald Trump: Fooled by your own politics. *Personality and Individual Differences*, 128, 21–24. <https://doi.org/10.1016/j.paid.2018.02.019>.
- Young, A., Ratner, K., & Fazio, R. (2014). Political attitudes bias the mental representation of a presidential candidate's face. *Psychological Science*, 25(2), 503–510. <https://doi.org/10.1177/0956797613510717>.