Sex Roles (2025) 91:44
https://doi.org/10.1007/s11199-025-01594-2

ORIGINAL RESEARCH

Standing Up to Sexism: Does Challenging Sexist Comments Have Transformational Benefits in How Women Value Themselves and Other Women?

Hazel Sayer[1] · **Marie Juanchich**[1] · **Veronica M. Lamarche**[1] · **Lilith A. Whiley**[1]

Accepted: 28 May 2025

Abstract

Research suggests that challenging sexist comments may mitigate its negative effects and may even have some positive effects, but the nature of this association is unclear: is it causal or a mere correlational artefact? Across three studies, we examined whether challenging sexist comments relates to how women value themselves and other women. We hypothesised that challenging sexist comments may have a transformative effect, a moderating effect, or no effect, merely reflecting pre-existing beliefs. Studies 1a, 1b, and 2 found that women who challenged sexist comments valued sexual consent more and accepted gender inequality and rape myths less. However, these effects disappeared when accounting for baseline sexism, suggesting that women with less sexist attitudes were more likely to challenge sexist comments – and already hold these values. Study 3 confirmed this, showing that challengers and those who wanted to challenge held similar attitudes, while non-challengers valued themselves and other women less. These findings question the assumption that challenging sexism has immediate transformative benefits – but the possibility remains that repeated acts of challenging would have stronger effects. Instead, our findings suggest that women who confront sexism are likely to have pre-existing egalitarian beliefs. For educators, activists, and policymakers, strategies to reduce sexist attitudes and fostering cultural and psychological change is paramount instead of simply placing the onus on women's individual acts of resistance.

Keyword Challenging sexism; Self-esteem; Rape myths; Sexual consent; Sex role attitudes

Sexism is a form of prejudice based on stereotypical beliefs about sex or gender (Dovidio et al., 2008) and functions to maintain patriarchy and gender hierarchies that disadvantage women. In more recent years, a spotlight has been shone on sexist culture and how it manifests in violence. The #MeToo, #TimesUp – and more recently in the United Kingdom, the #ReclaimTheseStreets social movements – were sobering wake-up calls, highlighting the prevalence of violence against women and girls as a result of a sexist and misogynistic culture. Further, sexism and misogyny have been found to permeate society, embedded in various sectors, with reports revealing sexism in healthcare (Topping, 2021), industry (for written evidence, see UK Parliament, 2024), education (National Education Union & UK Feminista, 2017), the entertainment industry (Women & Equalities Committee, 2024), policing (Baroness Casey of Blackstock DBE CB, 2023; Turner, 2024) and the justice system (Riverlight, 2024; The Fawcett Society, 2018).

Whilst everyday acts of sexism might seem innocuous, there are a range of negative personal consequences, such as detriment to self-esteem (Major et al., 2002; Swim et al., 2001) and propensity to self-blame following sexual assault (Camp, 2017) – as well as interpersonal consequences in how women are perceived, such as greater acceptance of systemic gender inequalities (Jost & Kay, 2005) and higher propensity to blame victims of rape and sexual assault (Chapleau et al., 2007). With sexist culture being at the root of so many negative and dangerous outcomes for women, reducing the pervasiveness of sexist attitudes is a global emergency and in line with the United Nations Development Goals (United Nations, 2024).

While the onus is on perpetrators of sexism to address their actions and behaviour in order to reduce sexist culture, there is research to suggest that challenging discrimination can act as a buffer against some of the physical and

✉ Hazel Sayer
hs18307@essex.ac.uk

[1] Department of Psychology, University of Essex, Essex, UK

Published online: 11 June 2025

psychological effects of experiencing discrimination for the targets (for review, see Chaney et al., 2015). These buffering effects include increased wellbeing (Foster, 2013; Sanchez et al., 2015), sense of empowerment and affecting change (Hyers, 2007), less reported distress in those that challenge (Noh & Kaspar, 2003), and less distress mediated by a sense of autonomy (Sanchez et al., 2015). There is also evidence that when women challenge sexism, they perceive themselves more positively. For example, challenging sexist comments has been found to increase competence, self-esteem, sense of empowerment (Gervais et al., 2010), greater wellbeing (Foster, 2013), and greater empowerment and wellbeing, with the caveat that the challenge must be deemed effective in changing the perpetrator's behaviour (Helwig, 2022). Challenging prejudice also possibly reduces stereotype use in perpetrators (Czopp et al., 2006) and observers (Rasinski & Czopp, 2010). Such promising findings have led to calls for further research on challenging sexism as an intervention for reducing sexist culture (Becker et al., 2014; Connor et al., 2017; Drury & Kaiser, 2014).

Women who hold attitudes more supportive of gender equality are more likely to identify sexism in sexist comments and therefore may be more likely to take self-protective actions by challenging it (Crocker & Major, 1989; Crocker et al., 1991; Landrine & Klonoff, 1997), therefore, are more likely to benefit from the positive personal outcomes associated with challenging. These findings suggest that women's support for gender equality may drive the positive personal outcomes associated with confronting sexism – meaning that what is often seen as a *consequence* of challenging sexism could, in fact, be partially caused be preexisting beliefs and attitudes. Certainly, these extant findings on the personal benefits of challenging sexism emphasise the need for more work on replicating personal benefits and extending the research to interpersonal benefits; not just how women see themselves, but also, how they perceive other women in the gender status quo. Therefore, this research aims to evaluate the effects of challenging sexism after exposure to sexist comments on a range of personal and interpersonal outcomes, specifically those that reflect the value women place on themselves and other women in society. This research will also explore whether there is a causal association between challenging sexism and these positive outcomes, or if the association is merely a correlational artefact.

The Personal and Interpersonal Negative Outcomes of Sexism for Women

Although everyday acts of sexism might appear harmless, research shows that these behaviours are linked to various negative impacts on how women perceive themselves, which we refer to here as 'personal outcomes.' Everyday sexism has been found to permeate society, such as in healthcare (Topping, 2021), industry (for written evidence, see UK Parliament, 2024), education (National Education Union & UK Feminista, 2017), the entertainment industry (Women & Equalities Committee, 2024), policing (Baroness Casey of Blackstock DBE CB, 2023; Turner, 2024) and the justice system (Riverlight, 2024; The Fawcett Society, 2018). Women may encounter comments that question their competence (e.g., "You're pretty smart for a woman") or employ gender stereotypes (e.g., "Calm down—women are so emotional"). Exposure to everyday acts of sexism through sexist comments affected women's wellbeing and their self-esteem (Swim et al., 2001), as well as inducing anxiety (Spencer et al., 1999), poorer performance on cognitive tasks (Dardenne et al., 2007) and decreased self-efficacy (when comments were perpetrated by a man; Jones et al., 2014). Women who feel that they have been discriminated against based on their gender have a higher prevalence of depression, increased psychological distress, poorer mental functioning, life satisfaction and self-rated health (Hackett et al., 2019). There is also evidence that sexist attitudes may relate to attitudes towards sexual consent, for example, holding negative attitudes towards women predicted greater rape myth acceptance, which affects how individuals attribute blame to rape and sexual assault survivors (Baugher et al., 2010). Further, holding sexist attitudes is related to the propensity to self-blame following sexual assault (Camp, 2017) as well as related to reducing rape to 'bad sex' (Peterson & Muehlenhard, 2004), which is associated with hesitancy in reporting rape and sexual offences to police (Lorenz et al., 2019).

Exposure to sexism might also shape how women view other women, referred to as 'interpersonal outcomes.' For example, reading sexist comments online strategically aimed at reaffirming traditional gendered norms and stereotypes, potentially further re-enforces the prevalence of gender inequalities (Felmlee et al., 2020). Research also indicates that exposure to sexist stereotypes reinforces negative attitudes towards women and prevents advances in their social status (Heilman, 2001; Romero-Sánchez et al., 2010; Tiggemann & McGill, 2004). Gender stereotypes that form the basis of sexism (e.g., women should be ladylike) are positively correlated with increased justification and acceptance of systemic gender inequalities in society (Jost & Kay, 2005). Research conducted across 51 countries showed that holding hostile sexist attitudes relates to acceptance of violence towards women (Herrero et al., 2017). Likewise, people with more sexist attitudes are more likely to blame women who are victims of rape and sexual assault and to exonerate the perpetrator—those who score high on hostile sexism are more likely to endorse 'rape myths,' such as the belief that women who report being raped might have played a part

in what happened, or that women lie about being raped to 'get back at men' (Chapleau et al., 2007), perpetuating rape myths that not just anyone can be a victim of rape – only those who bring it on themselves (Lonsway & Fitzgerald, 1994). Such myths make it difficult for victims to speak out, with less than 1 in 6 women in England and Wales who have been raped going on to report it to police (Office for National Statistics, 2020).

The extent of the consequences of receiving sexist comments might depend on the level of sexism that women themselves hold; if women reject sexism, the effects of experiencing sexism themselves may be more severe than in women who endorse sexism – which is not as unusual as one might expect. Large-scale studies involving up to 62 countries found that women endorse benevolent sexism as much as or more than men, including in the United Kingdom (Glick et al., 2000; Zawiska et al., 2025). Further, Bareket and Fiske (2023) synthesised multidisciplinary empirical literature on ambivalent sexism. Their findings suggest that both hostile and benevolent sexism persist as a coordinated system that sustains gender inequality. Hostile sexism manifests through envious and resentful prejudices, with an increased sensitivity to power dynamics and sexual cues, while benevolent sexism operates through prejudices related to interdependence, such as gender-based paternalism and role differentiation, reinforcing traditional gender roles and responding more strongly to role-related cues.

Benevolent sexism is more insidious and harmful than hostile sexism because it is often perceived as warmth, making it less likely to be recognised as sexist. In fact, men characterised as benevolently sexist are sometimes perceived as *more* supportive of gender equality. Benevolent sexism thus operates as a powerful tool that masks the ideological functions of sexism (Hopkins-Doyle et al., 2019). Moreover, research by Dardenne et al. (2007) shows that women exposed to benevolently sexist comments perform significantly worse on cognitive tasks, such as problem-solving, compared to those exposed to hostile sexism. Given these findings, our studies specifically focus on the more benevolent forms of sexism.

Research has explored why women may endorse sexism, particularly benevolent sexism, through the lens of system justification theory—Jost and Kay (2005) suggest that individuals have a psychological motivation to defend, justify, and rationalise the systems in which they live, even if these systems perpetuate inequality or disadvantage certain groups. When confronted with evidence of injustice, such as inequality or prejudice, individuals who strongly endorse belief in a just world (Lerner, 1980) may respond with defensive strategies that maintain their perception of fairness, often by denying, rationalising, justifying, or blaming victims rather than confronting the perpetrator (Hafer & Rubel, 2015). Expanding on ambivalent sexism theory, studies have linked feelings of powerlessness in an unequal society to greater endorsement of both benevolent and hostile sexism among women, as well as increased tolerance of gender inequality (Hammond et al., 2024). Particularly in the case of benevolent sexism, women are prepared to incur the costs of being 'undermined in competence, ambition and independence' for the security and protection that a chivalrous male partner will offer her. Further, highly anxious women in need of more relationship security rated men described as endorsing benevolent sexism as relatively more attractive and reported greater preferences for partners to hold benevolently sexist attitudes (Cross & Overall, 2018).

Furthermore, women are more likely to endorse benevolent sexism when thinking about traditional women, e.g., housewives, and hostile sexism when they are thinking about non-traditional women, e.g., feminists who fight against conforming to gender roles; career women (Becker, 2010). Women's endorsement of benevolent sexism can also lead to increased acceptance of hostile sexism over time (Sibley et al., 2007), leading women to internalise and perpetuate sexist ideologies, suggesting that benevolent sexism serves to justify and maintain existing gender hierarchies and traditional gender roles, where men are protectors/providers and women are dependent/submissive (Curun et al., 2017). This endorsement of sexism not only sustains the societal disadvantages of women (Glick & Fiske, 1996; Jost & Kay, 2005; Zawisza et al., 2025) but also reinforces harmful attitudes that justify violence against them (Zawisza et al., 2025). Given these consequences, a critical question arises: Can challenging sexism empower women and help mitigate its effects?

Can Challenging Sexist Comments Mitigate Their Negative Consequences?

Whilst the benefits of challenging sexist comments may seem advantageous to women, there are perceived costs of challenging sexism that could make it seem 'not worth it,' in an attempt to avoid being labelled as "cold," "lacking a sense of humour," and "oversensitive" (Becker et al., 2011). Likewise, women tend to avoid challenging sexism as they are concerned that they will seem impolite and aggressive (Swim & Hyers, 1999; Woodzicka & LaFrance, 2001) and can even adopt a strategy of silence in order to avoid further mistreatment or escalation (Ortiz, 2023). However, there has been a drive to motivate people to challenge sexism when they encounter it to reduce its prevalence (Becker et al., 2014; Council of Europe, 2000; Drury & Kaiser, 2014; Mallett & Wagner, 2011). Challenging prejudice (in terms of racism) has been found to reduce male perpetrators' use of sexist language (Mallett & Wagner, 2011), and confronting prejudice reduced the likelihood that perpetrators and observers would use negative stereotypes again in the future

(Czopp et al., 2006; Rasinski & Czopp, 2010), thus reducing the prevalence of prejudice in society. Furthermore, there is evidence that challenging sexist stereotypes motivates perpetrators to reduce subsequent racial stereotypes, which can have a ripple effect, reducing biases, thereby offering a promising strategy for promoting broader societal change and having a far-reaching effect over different stigmatised groups (Chaney et al., 2020).

Challenging sexism might also have an immediate personal benefit for the challengers. For example, challenging the perpetrator of a sexist comment in a staged online interaction was positively correlated with increased competence, self-esteem, and empowerment for women, relative to people who did not challenge it (Gervais et al., 2010). However, this study did not have a comparative non-sexist control condition, hence being correlational in nature, and difficult to ascertain whether the challenge itself influenced competence, self-esteem, and empowerment, or whether these are traits of people that challenge sexism. Furthermore, the 'challenge' was not in the form of verbally calling out the perpetrator in the heat of the moment, rather, participants rated the comment's appropriateness. The sample were also university students with a maximum age of 29 – research shows that younger women are more likely to subscribe to feminist views (Plan International, 2017).

Other benefits of challenging sexism include a feeling of restored control (Hyers, 2007), psychological wellbeing (Foster, 2013; Helwig, 2022), and less negative affect and greater wellbeing when forming part of a collective that reject sexist ideology on the internet (Foster, 2015; Foster et al., 2021). However, measuring the effects of challenging sexism has been somewhat mixed in findings; Swim and Hyers (1999) found no detrimental effect of sexist comments on self-esteem, nor any positive effect found when women challenged sexist comments (relative to a non-sexist control condition), and it was also acknowledged that women who are already committed to gender activism predicted the propensity to challenge sexist comments. Swim and Hyers (1999) also had a similar limitation as with Gervais et al.'s (2010) study described above, in that the opportunity to 'challenge' was more of an assessment of the comment than confrontation directed at the perpetrator. In response to the existence of contradictory findings, we conducted a series of experiments with larger, sample sizes and we followed open science practices with pre-registered method, hypotheses and analyses, and we publicly share data and analytical scripts on the Open Science Framework: https://osf.io/2pyv4

Present Research

Based on the current literature, three options are possible regarding the role of challenging sexism. First, it may be that challenging sexism has an effect of its own on women's value they hold for themselves and the value they hold for other women, that is additive to the effect of experiencing sexism. Second, challenging sexism might have a moderating effect that reduces the effect of experiencing sexism. Finally, it is possible that challenging sexism does not cause changes in women's self-value and value of other women – rather, being a challenger may relate to already possessing these specific psychological characteristics. Further research is thus needed to reconcile these possibilities with appropriate experimental and control conditions, more diverse samples, controlling for baseline sexist attitudes and avoiding the pressure of social responsibility by providing a private opportunity to challenge as opposed to a public one.

Study 1a

This study was preregistered on As Predicted (#46,153). In Study 1a, women received either a series of sexist comments or non-sexist comments (control condition) followed by an opportunity to challenge the comments, before measuring personal and interpersonal outcomes relating to the value women hold for themselves and other women. We compared the responses of women who challenged the sexist comments, those who did not challenge the sexist comments, and those who received neutral, non-sexist comments (control participants). This study extends prior research by including a control condition and by broadening the range of variables that may be affected by exposure to sexist comments as well as by challenging them. Past work mostly focused on personal outcomes for women (e.g., self-esteem, competence) but here, in addition to examining personal outcomes, we also examine effects on two variables related to how women perceive other women in the gender status quo: the tendency to justify systemic gender inequality and acceptance of rape myths.

We hypothesised that:

H1: Women exposed to sexist comments during a reasoning task will perform less well in that task and report lower self-esteem, place less value on sexual consent, and report greater justification for systemic gender inequality and rape myth acceptance compared to women who receive non-sexist comments.
H2: Women who challenge the sexist comments will have higher self-esteem, less justification for systemic gender inequality, and place higher value on sexual consent and lower rape myth acceptance, compared to those who do not challenge the sexist comments. We plan to explore how challengers and non-challengers fared relative to control participants.
H3: Sexist attitudes will be negatively related to participants' propensity to challenge the sexist comments, posi-

tively related to justification of systemic gender inequality and rape myth acceptance, and negatively related to self-esteem and the value placed on sexual consent.

Method

Participants and Procedure

Participants were recruited to the study via the Prolific platform and were compensated £2.00 for a median completion time of 16 minutes (an average of £8.00 per hour). They were unaware that the study was aimed at only women. Participants were told that they were taking part in a study investigating how they see themselves and others and how it impacts performance on a cognitive task to avoid revealing the full nature of the study. Following informed consent, participants were asked for their age and gender. If participants identified as a man, non-binary or other, they received a notification that the study was aimed at those that identified as women, with the option to either opt out or continue for inclusivity (but data of those who identified as men were filtered out for the analyses).

Five hundred and forty-one participants took part in the study. We filtered participants to include those who registered on the Prolific platform as female, but also had to remove 13 cases where participants did not identify as women, but we retained data from two participants who identified as non-binary and were given the choice to continue given that the study was aimed at women. The analytical sample therefore included 529 participants which allowed capture of a medium effect size of the sexism manipulation on the dependent variables (Cohen's $d = .056$, alpha = .05 and power = .90). Participants were between the ages of 18 and 82 ($M_{age} = 34$, $SD = 13.12$). Participants were White British ($n = 179$, 33.8%), White other ($n = 253$, 47.8%), Black British ($n = 9$, 1.7%), Black other ($n = 29$, 5.5%), Asian British ($n = 17$, 3.2%), Asian other ($n = 25$, 4.7%), and other ethnicity ($n = 17$, 3.2%).

After the consent stage, participants were 'introduced' to a male character (John), who portrayed himself as the study lead, via audio instructions recorded for each of the blocks of the study. This feature of the instructions created the impression that subsequent comments were made by a man, considering past research showing decreased self-efficacy only when sexism was perpetrated by men as opposed to other women (Jones et al., 2014). Participants first completed a measure of sexism, followed by a cognitive test that featured the sexism manipulation (sexist comment vs. non-sexist control), and then the opportunity to respond to the comments. Participants were randomly allocated to the sexist or non-sexist conditions at a ratio of 77:23, respectively, so that we could compare the control participants (23%) to two roughly equivalent groups within the sexist condition: those who challenged the comments and those who did not. We decided on this ratio as we expected around half of participants to challenge the sexist comments, and we were most interested in comparing those that did challenge and those that did not within condition. Such a ratio and expected confrontation rates would provide a sample of 416 participants split into two groups of 208, giving a 90% power to detect a small-medium effect of confrontation (versus no confrontation) on the outcome variables, $d = 0.29$. Finally, participants completed measures of self-esteem, justification of gender inequality, value placed on sexual consent and rape myth acceptance (randomly presented to participants) before completing some sociodemographic questions and being debriefed.

Materials

Sexist Attitudes

All participants completed the Ambivalent Sexism Inventory (Glick & Fiske, 1996), comprising 22 items with two subscales of coexisting ideologies: Hostile Sexism, a sample item is "Women seek to gain power by getting control over men" and Benevolent Sexism, a sample item is "Women should be cherished and protected by men," $\alpha = .88$. Participants rated their agreement with the statements using a 6-point scale ranging from 0 (*disagree strongly*) to 5 (*agree strongly*). We created a sexist attitude index using each participant's average score, of which higher scores indicate stronger sexist attitudes.

Verbal Cognitive Reflection Test

The virtual research assistant 'John' then introduced the verbal reasoning task, which was six items taken from the verbal Cognitive Reflection Test (Sirota & Juanchich, 2018), hereafter referred to as the CRT-V. To avoid the confounding effects on self-esteem of having difficulty with the questions, we added four multiple choice responses to each question to simplify the task. As sample item is "Mary's father has five daughters but no sons – Nana, Nene, Nini, Nono. What is the fifth daughter's name? A) There's no way of telling B) Nunu C) Mary D) Nano." Correct answers were summed to give each participant a performance score from a range of 0 to 6, with a score of 6 representing correct answers on all items. The average performance for this task was 3.40 ($SD = 1.85$).

Sexism Manipulation

The sexist (or control) comments were embedded within the CRT-V (see Table 1). We chose a series of sexist comments that could be characterised as patronising and infantilising,

Table 1 Sexist comments versus non-sexist comments embedded in the CRT-V

Comment delivery	Non-sexist version	Sexist version
Introduction comment	Participants tend to find verbal reasoning tasks quite tricky, so to assist you, I have included multiple choices, which should help	**Women** tend to find verbal reasoning tasks quite tricky, so to assist you, I have included multiple choices, which should help
Interim comment A	Keep going! Try to stay focused. **Participants** can let their emotion get in the way	Keep going! Try to stay focused. **Girls** can let their emotion get in the way
Interim comment B	Only three more to go!	Only three more to go **sweetie**!
Concluding comment	You're all done! Good work, **participants** don't usually do well on that task!	You're all done! Good work, **girls** don't usually do well on that task!"

Note. Text in bold is to highlight distinction between sexist and neutral comments and were not presented in bold to participants.

similar to those used in past research (e.g., Lamarche et al., 2020). Although these kinds of comments do not fall under the hostile categorisation of sexism, they are still undermining women's sense of competence and autonomy regardless of their alleged 'benevolence.' These comments were pretested on a sample of 109 women who rated sexist versus non-sexist comments. The pretest confirmed that the sexist comments were more undermining, more sexist, and less supportive than the control comments. The data and full details of the development of these sexist comments and the pretest are available in supplementary materials on Open Science Framework: https://osf.io/2pyv4/ (see SM1).

Challenging Sexist Comments

After completing the CRT-V, participants were provided with an opportunity to leave feedback on the instructions that 'John' gave them throughout the task, which provided participants with the chance to challenge the sexist comments in their own words in a free text box, presented as follows:

> *"Hi, John here. Thank you for completing the task. Do you have any comments on the instructions given to you throughout this task? If you do not have any comments for me, please state 'none' in the box."*

Participants' written reactions to the comments were coded by the first author. Coding participants' challenges or lack thereof were based on a coding framework similar to Swim and Hyers (1999), where two female coders independently noted any challenge deemed as verbal expressions of displeasure or disagreement with the sexist comments. Responses that were considered challenges in Swim and Hyers' framework included polite resistance, such as requesting clarification or confusion about the perpetrator's intent, e.g., "What did you say?" or "What do you mean?" as well as use of sarcastic humour as part of their challenge responses. We took the same stance in our own coding framework in that challenges did not have to be risky or overtly confrontational to be coded as a challenge, but they needed demonstrate resistance by saying *something* considered to be a clear expression of displeasure or disagreement.

Responses were coded as follows: 0: = Unclear/no challenge*;* 1 = Positive reaction (e.g., *"Really clear instructions and I enjoy hearing the voice of a real human giving these instructions*"); 2 = Expression of confusion (e.g., *"that was very strange commentary!*); 3 = Respond with sarcastic joke (e.g., "*You did a good job typing out all those questions. Men aren't usually good at verbal tasks—but you did great!"*; 4 = Requesting clarification (e.g., "*Where are you going with this*?"); 5 = Expressing negative feeling (e.g., "*The comments were very insulting*"); 6 = Points out sexism (e.g.,"*Patronising, sexist remarks!*"); 7 = Explanatory challenge (e.g., "*Constant reminders telling me that I'm a girl and women tend to do worse—felt condescending. I didn't like that*"); 8 = Direct challenge (e.g., "*Do NOT stereotype women like that").* These responses were further coded into superordinate categories of *did not challenge* (responses 0 and 1) and *challenged* (responses 2–8).

Before answering the sociodemographic questions, all participants had their responses about the comments from John on the CRT-V presented back to them and they were asked what their intention was in their response to 'John,' before being debriefed. We evaluated the validity of our coding scheme by comparing whether an answer was coded as challenging or not and the intention behind the response that participants self-reported at the end of the study (e.g., "I tried to express my negative impression to challenge the person running the study."; "I tried to express my positive impression."; "I did not have any particular opinion."). The consistency between self-reported intentions and author coding of responses as challenging was good, indicating that the intentions of participants were explicit in their responses (Cohen's $\kappa = .73$). For example, of the answers where participants reported that they tried to express their negative opinion, 74% were coded as challenging. Following the challenge opportunity, all participants completed four measures encapsulating their perceived value of themselves and other women, or what we refer to as personal and interpersonal

outcomes, which were presented on separate pages and in a randomised order.

Self-Esteem

The Rosenberg Self-Esteem Scale (Rosenberg, 1965; α = .85) was used to assess level of self-esteem. This is a 10 item, 4-point Likert scale, with items rated from 0 (*Strongly disagree*) to 3 (*Strongly agree*). A sample item is "I feel that I have a number of good qualities." We created a self-esteem index using each participant's mean score, where higher scores indicate greater self-esteem.

Value Placed on Sexual Consent

We assessed how much value women placed on sexual consent using the 'positive attitude toward establishing consent' subscale from the Sexual Consent Scale (revised)—Subscale 2 (Humphreys & Brousseau, 2010). This scale consisted of 11 items, with a 7-point Likert scale ranging from 0 (*Strongly disagree*) to 6 (*Strongly agree*). A sample item is "I feel it is the responsibility of both partners to make sure sexual consent is established before sexual activity begins" (α = .91). We computed a value placed on sexual consent index using each participant's mean score, where higher scores indicate greater value placed on sexual consent.

Justification of Gender Inequality (JGI)

The Gender-Specific System Justification Scale (Jost & Kay, 2005, adapted from Jost & Banaji, 1994) assessed support for the gender status quo, and consists of eight items scored on a 7-point Likert scale, from 1 (*strongly disagree*) to 7 (*strongly agree*). A sample item is "Most policies relating to gender and the sexual division of labour serve the greater good" (α = .71). We created a gender-specific system justification index using each participant's mean score, where higher scores indicate greater justification of gender inequality.

Rape Myth Acceptance

We used the Illinois Rape Myth Acceptance Scale short-form (Payne et al., 1999) to assess endorsement of rape myths, which comprises 20 items and we used a 5-point Likert scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). A sample item is "If a woman is raped while she is drunk, she is at least somewhat responsible for letting things get out of control" (α = .93). We created a rape myth acceptance index using each participant's mean score, where higher scores indicate greater rape myth acceptance.

Results and Discussion

Effect of Exposure to Sexist Comments

We used a MANOVA to test the effect of exposure to sexist comments (exposure versus control) on personal and interpersonal consequences for women. In contrast with our expectation, we did not find evidence that exposure to sexist comments (versus control) had an effect on performance in the verbal reasoning task, self-esteem, value of sexual consent justification of gender inequality, or rape myth acceptance, respectively, $F(1,527) = 0.01$, $p = .945$, $\eta_p^2 < .001$; $F(1,527) = 1.30$, $p = .254$, $\eta_p^2 = .002$; $F(1,527) = 0.54$, $p = .463$, $\eta_p^2 = .001$; $F(1,527) = 1.57$, $p = .210$, $\eta_p^2 = .003$; $F(1,527) = 0.04$, $p = .952$, $\eta_p^2 < .001$.

Link Between Challenging Sexist Comments and Women's Value of Self and Other Women

While exposure to sexism did not have the expected effect on women's perceptions of themselves and other women, many women did spot the sexism in the comments and reacted to it. In the sexist condition ($n = 401$), 49% of participants challenged the sexist comments (e.g., "Do NOT stereotype women like that" and "[the] comments were patronising and sexist towards women," see Table 2). To test the link between challenging sexist comments and how women perceive themselves and other women, we conducted a MANOVA comparing women who chose to challenge the sexist comments, women who chose not to challenge, and women who were not exposed to any sexist comments and included the four outcome variables as dependent variables (see Table 3).

As shown in Fig. 1, the results partly support the hypothesis (H2) that women who challenge the sexist comments will have less justification for systemic gender inequality, place higher value on sexual consent, and report lower rape myth acceptance, compared to women that do not challenge the sexist comments. The data did not support the expectation that women who challenge the

Table 2 Study 1a, Responses to comments in sexist versus non-sexist condition

Response	Sexist Condition ($n = 401$)	Non-Sexist Condition ($n = 128$)
No challenge/unclear response	$n = 190$, 47%	$n = 115$, 90%
No challenge: Positive reaction	$n = 14$, 4%	$n = 13$, 10%
Challenged	$n = 197$, 49%	$n = 0$, 0%

Note. $N = 529$.

sexist comments have higher self-esteem, compared with women who do not challenge. Women who did not challenge the sexist comments reported a higher level of self-esteem, compared with those that did challenge the sexist comments, who reported similar levels of self-esteem to women in the control condition.

We tested the pairwise comparisons using Bonferroni correction between challenge versus no challenge, challenge versus control, and no challenge versus control. As shown in Table 4, women who did challenge the sexist comments had similar levels of self-esteem than those who did not challenge them and those in the control condition. As hypothesised, those that challenged the comments had statistically significantly lower levels of justification for gender inequality, placed more value on sexual consent, and were less accepting of rape myths compared to non-challengers. Challengers also had statistically significantly lower justification for systemic gender inequality and lower rape myth acceptance compared to those in the non-sexist control condition.

Table 3 Mean (*SD*) of personal/interpersonal outcomes for women as a function of their response to sexism compared to non-sexist control condition

Outcome	Sexism Overall ($N=529$)	Sexism Challenged ($n=202$)	Sexism No Challenge ($n=199$)	Control Condition ($n=128$)	F (2, 526)	p	η_p^2
Self-esteem	1.99 (0.51)	1.96 (0.50)	2.05 (0.55)	1.95 (0.45)	2.28	.103	.009
Value of sexual consent	5.76 (0.94)	5.89 (0.87)	5.66 (1.00)	5.71 (0.93)	6.33	.002	.023
Justification of gender inequality	3.71 (0.96)	3.48 (0.96)	3.90 (0.87)	3.81 (1.01)	10.66***	< .001	.039
Rape myth acceptance	1.49 (0.61)	1.32 (0.39)	1.68 (0.69)	1.49 (0.61)	19.81***	< .001	.070

Note. *** $p < .001$, ** $p < .01$, * $p < .05$.

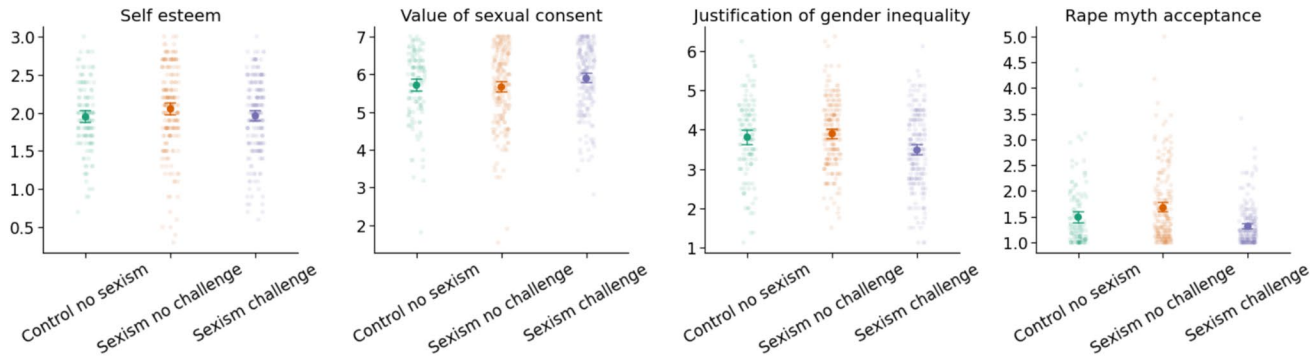

Fig. 1 Mean Scores of personal/interpersonal outcomes for women as a function of challenging compared to control condition. *Note.* The preregistered analyses included first a comparison of the challengers and non-challengers in the sexist condition only, followed by a second ANOVA testing how those two groups related to the control condition. For completeness, we conducted an ANOVA that compared the challenge vs. no challenge in the sexist condition and included the non-sexist control condition. An ANOVA comparing challenge vs. no challenge in the sexist condition only can be found in SM3.

Table 4 Pairwise comparisons for the effect of challenging sexist comments on personal and interpersonal outcomes for women

Outcome	Challenge vs. No Challenge			Challenge vs. Control			No Challenge vs. Control		
	Mdiff	CI	*p*	*Mdiff*	CI	*p*	*Mdiff*	CI	*p*
Personal									
Self-esteem	-.09	-.21, .03	.215	.01	-.12, -.15	> .999	.10	-.03, .24	.205
Value of sexual consent	.23	-.01, -.46	.041*	.18	-.07, .44	.243	-.05	-.30, .21	> .999
Interpersonal									
Justification of gender inequality	-.42	-.64, -.19	< .001***	-.33	-.58,-.07	.006**	.09	-.17, .34	> .999
Rape myth acceptance	-.36	- .50, -.22	< .001***	-.18	-.33, -.02	.014*	.19	-.03, .34	.013*

Note. *** $p < .001$, ** $p < .01$, * $p < .05$.

Whether women challenged the sexist comments may have also been partially explained by participants' sexist attitudes. A logistic regression confirmed that sexist attitude was indeed statistically significantly associated with the propensity to challenge sexism, with women scoring higher in sexist attitude challenging the comments less often, χ^2 (1) = 45.02 $p < .001$, Cox & Snell $R^2 = .106$, β = 1.16, Wald (1) = 39.20, $p < .001$. Sexist attitude was also positively correlated with self-esteem, $r = .17$, $p < .001$, justification for systemic gender inequality, $r = .49$, $p < .001$ and rape myth acceptance, $r = .63$, $p < .001$, and negatively correlated with value placed on sexual consent, $r = -.29$, $p < .001$.

To evaluate whether baseline sexist attitudes could explain the link between challenging sexist comments and personal and interpersonal outcomes, we ran a MANCOVA, including baseline sexism as a covariate to control for sexist attitudes in our model. When baseline sexism was included as a covariate, the effects on all dependent variables became non-significant.: self-esteem F (3, 525) = 2.65, $p = .348$, $\eta_p^2 = .004$, value of sexual consent, F (3, 525) = 1.07, $p = .899$, $\eta_p^2 < .001$, support for systemic gender inequality, F (3,525) = 0.45, $p = .635$, $\eta_p^2 = .002$, and rape myth acceptance, F (3,525) = 1.66, $p = .191$, $\eta_p^2 = .006$.

The MANCOVA results indicate that whilst we observe some differences between those who challenged sexist comments and those who did not, the differences were mostly explained by participants' initial level of sexism. Furthermore, unexpectedly, not challenging was associated with greater self-esteem, even after we controlled for baseline sexism levels. This might be the case because the warm undertone of the sexist comments could have boosted self-esteem in women who were unbothered by the sexist nature of the comments (and hence did not challenge). Alternatively, those with higher self-esteem might not have felt the need to address the sexist comments.

Yet another unexpected result was the null effect of the sexist comments on all dependent variables, which does not support previous findings in the literature, though is consistent with Swim and Hyers (1999), who also failed to find an effect of the sexist comments. In our case, we considered how this might be explained by the study design: the presence of the sexism scale at the onset of the study may have forewarned participants about the topic under investigation and made the subsequent sexist comments appear less genuine. To evaluate this possibility, we extended Study 1 by inviting women from the control condition (and for whom we had already assessed sexist attitude levels) to take part in a new study four months later, where they were exposed to the sexist comments without the sexist attitude measures.

Study 1b

This study was preregistered on As Predicted (#51,842). This study follows on from Study 1a by re-testing participants from the control condition in the sexist condition, without the sexism measure at the onset of the study. We did this to avoid forewarning participants about the goal of the study, while still having the baseline measure of sexism from Study 1a (Time 1). The design left enough time between the two experimental conditions for participants not to remember their previous answers and give the possibility to control for initial sexist attitudes. We therefore had the same hypotheses in Study 1a; however, we used a repeated measures design to compare participants at Time 1 (T1) in the no sexist condition to the same participants at Time 2 (T2) in the sexist condition.

H1: Women will perform less well on the reasoning task, report lower self-esteem, place less value on sexual consent, and express greater justification for systemic gender inequality and rape myth acceptance in the sexist comment condition at T2 compared to the non-sexist condition at T1.
H2: We expect that women who challenge the sexist comments will have higher self-esteem, less justification for systemic gender inequality, place higher value on sexual consent and lower rape myth acceptance, compared to those who do not challenge the sexist comments.
H3: Sexist attitudes measured at T1 will be negatively related to participants' propensity to challenge the sexist comments, and sexist attitudes will be positively related to justification of systemic gender inequality and rape myth acceptance, and negatively related to self-esteem and value placed on sexual consent.

Method

Participants

The study was advertised to the 128 participants that took part in the control condition in Study 1a four months prior on the Prolific platform for one week. Of the 98 participants who had first clicked on the study, 91 completed fully (though nine participants opted to bypass the rape myth acceptance measure), leaving an analytical sample of 91 female participants aged 18–72 years, $M_{age} = 36.30$, $SD = 13.98$. A sensitivity analysis was conducted to capture a small to medium effect size of challenging sexism on the dependent variables, $F = 0.20$, alpha = .05 and power = 0.90. Participants were White British ($n = 179$, 33.8%), White other ($n = 253$, 47.8%), other ethnic

background ($n = 17$, 3.2%), Asian other ($n = 25$, 4.7%), Asian British ($n = 17$, 3.2%), Black other ($n = 29$, 5.5%) and Black British ($n = 9$, 1.7%).

Procedure

We used a one-way repeated-measures design (T1 No-sexism vs. T2 Sexism), though as before, we treat the challenges to sexism responses (challenged comments versus did not challenge comments) as an independent variable in our analyses. Participants were paid £1.60 to take part, with an average completion time of 12.38 minutes, therefore participants were paid an average of £7.75 per hour. The retest study was the same in materials and procedure as the sexist condition of Study 1a, except that participants did not answer the Ambivalent Sexism Inventory at the onset of the study. The challenge to sexism response was also, once again, coded by the first author using the same coding framework as in Study 1a, and validated using participant's self-identified intention in their response to 'John.' The consistency between author coding of the challenge response and self-reported intention was excellent, indicating that the intention of participants was explicit in their responses (Cohen's $\kappa = .86$). For example, of the answers where participants reported that they tried to express their negative opinion, 92% were coded as challenging.

Results and Discussion

The Effects of Exposure to Sexism

We hypothesised that being exposed to sexist comments would result in poorer performance in the CRT-V, lower self-esteem, higher justification of gender inequality, less value placed on sexual consent, and higher rape myth acceptance. We conducted a one-way repeated-measures MANOVA to examine the effect of exposure to sexism (T1: No-sexism versus T2: Sexism) on CRT-V performance, self-esteem, justification of gender inequality, and rape myth acceptance. Contrary to our expectations and consistent with results of Study 1a, being on the receiving end of sexism had no effect on verbal reasoning performance, self-esteem, value of sexual consent, and rape myth acceptance (see Table 5). The sexist manipulation only had a statistically significant effect on the justification for gender inequality, but in the opposite direction than expected.

The Effects of Challenging Sexism

Despite the sexist comment not having the expected effect on the outcome measures, we tested the effect of challenging sexism at T2 on the outcome measures. Forty-three percent of participants challenged sexism, similar to the challenge rates of those in the sexist condition in Study 1a (see Table 6). To test the link between challenging sexist comments and how women perceive themselves and other women, we conducted a MANOVA comparing women who chose to challenge the sexist comments and women who chose not to challenge and included the four outcome variables as dependent variables.

As expected, there was a statistically significant association with value placed on sexual consent, justification of gender inequality, and rape myth acceptance; however, we did not observe a statistically significant association between challenging and self-esteem (see Table 7).

Table 6 Study 1b, responses to comments in sexist condition at T2

Challenge response	Proportion of participants
No challenge / unclear response	$n = 51$, 56%
No challenge: Positive reaction	$n = 1$, 1%
Challenged	$n = 39$, 43%

Note. $N = 91$.

Table 5 Mean (*SD*) for main study variables as a function of exposure to sexism

Outcome	Control (T1)	Sexist (T2)	F	p	η_p^2
CRT-V	3.31 (1.81)	3.62 (1.91)	2.31	.132	.025
Personal					
Self-esteem	1.92 (0.46)	1.98 (0.46)	2.48	.119	.027
Value of sexual consent	5.75 (0.95)	5.84 (0.91)	2.16	.145	.023
Interpersonal					
Justification of gender inequality	3.77 (1.04)	3.59 (0.99)	6.95*	.010	.072
Rape myth acceptance	1.43 (0.48)	1.41 (0.50)	0.32	.574	.004

Note. $N = 91$ participants were re-tested; however, nine chose not to complete the RMA scale, in which case, $n = 82$ for RMA. $df = 1, 90$, except for rape myth acceptance, $df = 1, 81$.

*** $p < .001$, ** $p < .01$, * $p < .05$.

Table 7 Mean (*SD*) of the personal/interpersonal outcomes for women as a function of their response to sexism (challenge or no challenge) at T2

Outcome	Challenged ($n=38$)	No Challenge ($n=53$)	F	p	η_p^2
Personal					
Self-esteem	1.89 (0.53)	2.04 (0.45)	2.00	.161	.022
Value of sexual consent	6.22 (0.67)	5.56 (0.95)	13.55***	< .001	.132
Interpersonal					
Justification of gender inequality	2.98 (0.93)	4.03 (0.77)	34.24***	< .001	.278
Rape myth acceptance (*df* 1, 81)	1.17 (0.20)	1.67 (0.57)	15.22***	< .001	.160

Note. $N=91$ and $N=82$ for rape myth acceptance scale. *** $p<.001$, ** $p<.01$, * $p<.05$.

As with Study 1a, the results partly support the hypothesis (H2) that women who challenge the sexist comments will have less justification for systemic gender inequality, place higher value on sexual consent and report and report lower rape myth acceptance, compared to women that do not challenge the sexist comments. The data did not support the expectation that women who challenge the sexist comments have higher self-esteem, compared with women who do not challenge.

To evaluate whether baseline sexist attitudes could explain the link between challenging sexist comments and personal and interpersonal outcomes, we ran a MANCOVA, including baseline sexism as a covariate to control for sexist attitudes in our model (measured at T1), When baseline sexism was included as a covariate, the effects on all dependent variables became non-significant: self-esteem $F(1,91)=0.31$, $p=.577$, $\eta_p^2=.004$, value of sexual consent, $F(1, 91)=0.35$, $p=.553$, $\eta_p^2=.004$, support for systemic gender inequality, $F(1,91)=2.26$, $p=.136$, $\eta_p^2=.025$ and rape myth acceptance, and $F(1,82)=1.19$, $p=.279$, $\eta_p^2=.015$.

A logistic regression confirmed that sexist attitude (measured at T1) was statistically significantly associated with the propensity to challenge sexism at T2, with women scoring higher in sexist attitude challenging less often, $\chi^2(1)=35.98$, $p<.001$, Cox & Snell $R^2=.327$, $\beta=-.287$, Wald $(1)=12.97$, $p<.001$. Sexist attitudes were also positively correlated with justification for systemic gender inequality, $r=.60$, $p<.001$, negatively correlated with value placed on sexual consent, $r=-.45$, $p<.001$ and positively correlated with rape myth acceptance, $r=.54$, $p<.001$, but sexist attitudes were not correlated with self-esteem, $r=.17$, $p=.107$.

Data from Studies 1a and 1b do not support the expectation that receiving sexist comments has immediate detrimental outcomes in women's personal value and the value they hold for other women. However, we found correlational evidence of the link between sexist attitudes and the value women place on sexual consent, their justification for gender inequality, and rape myth acceptance. These associations became null when we controlled for baseline sexism. Finally, when we compared women who challenged sexism and those who did not, we found expected differences in value placed on sexual consent, justification for gender inequality, and rape myth acceptance.

Study 2

We preregistered this study on As Predicted (#54,641). Study 2 retests the first two hypotheses from Study 1a, in which we compared women's value of themselves and other women, depending on whether they received sexist or non-sexist comments, and whether they chose to challenge sexist comments or not. However, we observed feedback in Study 1a that indicated that the presence of the Ambivalent Sexism Inventory at the onset signposted the goal of the study. To mitigate this, Study 1b retested participants four months later without the presence of the Ambivalent Sexism Inventory, though on a smaller sample size due to retesting those from the control condition of Study 1a. We found consistent results with Study 1a that receiving sexist comments did not have the expected immediate effect on women's perceptions of themselves and other women, but that challenging sexist comments were related to better outcomes in terms of higher value placed on sexual consent, lower justification of systemic gender inequality and lower rape myth acceptance, compared to non-challengers. These associations became null when controlling for sexist attitudes. However, we wanted to conduct Study 2 without the Ambivalent Sexism Inventory that measures sexist attitudes to avoid signposting the goal of the study before exposing participants to sexist comments with a larger sample to increase statistical power and our ability to detect an effect of exposure to sexism and differences related to challenging.

Method

Participants

The study was advertised using a snowball sampling method on social media over a period of four weeks. From an initial 744 participants who clicked on the survey link, 383 were

removed for < 80% completion. We decided to this lower completion threshold (relative to Study1a and 1b) because of a larger number of participants who did not complete the study fully. We therefore included all participants that completed up to the sociodemographic section at the end of the study (up to 80%). Six participants did not complete the consent form and were hence removed from the dataset, and 29 participants were removed for not meeting the gender requirements for this study. This left a total sample of 326, to capture a medium effect size of challenging on dependent variables, $F = 0.04$, alpha = .05 and power = 0.90. Three hundred and twenty-four participants identified as women, with one participant identifying as non-binary and one stated they would prefer not to disclose their gender but opted to continue with the study. Participants were aged 18 -71, $M = 28$, $SD = 11.98$. Participants were White other ($n = 107$, 32.8%), White British ($n = 102$, 31.6%), Black British ($n = 40$, 12.4%), Asian/other ($n = 23$, 7.1%), Black/other ($n = 19$, 5.1%), other ethnic background ($n = 19$, 5.9%), Asian British ($n = 5$, 1.5%) and mixed British/mixed other ($n = 8$, 2.4%).

Procedure

Study 2 aimed to replicate Study 1a, except that we did not include the Ambivalent Sexism Inventory. The allocation of participants to the experimental and control conditions was again 77:23. Once again, as with the previous studies, participant reactions to the comments were coded by the first author, using the same coding schema as in Studies 1a and 1b. Once again, we evaluated the validity of the coded responses with participants' self-report, where we asked participants what their intentions were with their reaction (e.g., "I tried to express my negative impression to challenge the person running the study"; "I tried to express my positive impression"; "I did not have any particular opinion"). The consistency between self-reported intentions and author coding was good, indicating that the intentions of participants were explicit in their responses (Cohen's κ = .78). For example, of the answers where participants reported that they tried to express their negative opinion, 82% were coded as challenging.

Results and Discussion

The Effect of Exposure to Sexism and Challenging Versus Control

Once again, we used a MANOVA to test the effect of exposure to sexism versus no sexism on women's value of themselves and other women. The MANOVA revealed that there was no statistically significant effect of exposure to sexism on any of the dependent variables (see Table 8). We further tested the effect of challenging sexism compared to not challenging sexism and the control condition. In the sexist condition ($n = 258$), 44% of participants challenged the sexist comments ($n = 113$), which is consistent with our findings from Studies 1a and 1b (see Table 9). To test the link between challenging sexist comments and how women perceive themselves and other women, we conducted a MANOVA comparing women who chose to challenge the sexist comments and women who chose not to challenge and included the four outcome variables as dependent variables.

As expected, there was a statistically significant association between challenging and value placed on sexual

Table 9 Study 2 responses to comments in sexist condition versus non-sexist condition

Challenge response	Sexist Condition ($n = 258$)	Non-Sexist Condition ($n = 68$)
No challenge / unclear response	$n = 137$, 53%	$n = 67$, 99%
No challenge: Positive reaction	$n = 8$, 3%	$n = 1$, 1%
Challenged	$n = 113$, 44%	$n = 0$, 0%

Note. $N = 326$.

Table 8 Mean (*SD*) of the personal/interpersonal outcomes for women as a function of exposure to sexist versus non-sexist comments

Outcome	*Control*	*Sexist*	*F*	*p*	η_p^2
Verbal cognitive reflection test	3.79 (1.86)	3.59 (1.69)	0.65	.420	.002
Personal					
Self-esteem	1.92 (0.43)	1.91 (0.47)	0.002	.960	< .001
Value of sexual consent	5.84 (0.95)	5.92 (0.85)	0.38	.540	.001
Interpersonal					
Justification of gender inequality	3.32 (0.86)	3.28 (0.89)	0.08	.777	< .001
Rape myth acceptance	1.31 (0.31)	1.34 (0.40)	0.16	.260	.001

Note. $df = 1, 324$, except for RMA because some people chose to skip this scale, $df = 1, 309$. In Study 1, we used a 9-point Likert for Justification of Gender Inequality, which was reduced to a 7-point scale in Study 2 onwards. This was because it was too wide for participants to see in full on Qualtrics, which meant the upper end was cut off. *** $p < .001$, ** $p < .01$, * $p < .05$.

Table 10 Mean (*SD*) of the personal/interpersonal outcomes for women as a function of their response to sexism compared to a non-sexist control condition

Outcome	Sexism Overall ($N=326$)	Sexism Challenged ($n=113$)	Sexism No Challenge ($n=145$)	Control Condition ($n=68$)	*F*	*p*	η_p^2
Self-esteem	1.92 (0.46)	1.86 (0.49)	1.96 (0.45)	1.92 (0.43)	1.56	.212	.010
Value of sexual consent	5.90 (0.85)	6.01 (0.76)	5.84 (0.86)	5.84 (0.95)	1.47	.232	.009
Justification of gender inequality	3.29 (0.89)	3.07 (0.89)	3.45 (0.86)	3.32 (0.86)	6.26**	.002	.037
Rape myth acceptance	1.33 (0.38)	1.26 (0.31)	1.40 (0.45)	1.31 (0.31)	4.22**	.016	.027

Note. $df=2, 323$, except for RMA because some people chose to skip this scale, $df=2, 308$.
*** $p<.001$, ** $p<.01$, * $p<.05$.

consent, justification of gender inequality, and rape myth acceptance; however, we did not observe a statistically significant association between challenging and self-esteem (see Table 10). Pairwise comparisons using Bonferroni correction between challenging versus not challenging the sexist comment versus the control condition showed that there was a statistically significant difference between challenging versus not challenging on justification of gender inequality and rape myth acceptance, but no difference relative to the control condition (see Table 11).

Consistent with Study 1a and 1b's results, the findings from Study 2 indicate that exposure to sexist comments did not alter women's value of themselves and other women. Furthermore, we found that while we observe some differences between those who challenge sexism and those who did not in terms of how women perceive other women, there was no difference relative to the control (no sexist comment) group. The results suggest that women's value for other women was not due to challenging the sexist comments, but rather the driver of challenging the comments.

Study 3

This study was preregistered on As Predicted: https://aspredicted.org/VPV_HZR.

To disentangle the act of challenging sexist comments from the psychological attributes of the person challenging the comments, we conducted a conceptually similar study with an alternative experimental design and a different set of sexist comments. In Studies 1a, 1b, and 2, all the participants were given the possibility to challenge the sexist comment. We compared participants who challenged and those who did not challenge the sexist comments prior to answering the outcome measures, whilst controlling for sexist attitudes in Studies 1a and 1b. In Study 3, we include an additional comparison group: women who want to challenge but were not given the opportunity to do so before answering the outcome measures. Challenging is posited to interrupt the rumination process following exposure to sexist comments (Hershcovis et al., 2018), implying that those who challenge sexism will report better outcomes than those who ruminate following exposure. In this experimental design, all the participants were exposed to sexist comments. Seventy percent of participants were given the possibility to challenge the sexist comments and then complete the outcome variable measures. The remaining 30% of participants were not given the possibility to challenge until the end of the study. Hence, we can differentiate between being the type of person that challenges compared to those who do not, and how both might affect women's value of themselves and other women relative to not challenging and not being a challenger.

This study also provides the opportunity to examine the role of baseline sexist attitudes in challengers, specifically those who did challenge immediately, and those who were not provided with an opportunity to do so until after they had completed outcome measures, but still opted to challenge. We therefore reintroduced the Ambivalent Sexism Inventory

Table 11 Pairwise comparisons for the effect of challenging sexist comments on personal and interpersonal outcomes for women

Outcome	Challenge vs. no challenge			Challenge vs. control			No challenge vs. control		
	Mdiff	*CI*	*p*	*Mdiff*	*CI*	*p*	*Mdiff*	*CI*	*p*
Self-esteem	-.10	-.24, .04	.236	.06	-.23, .11	> .999	-.04	-.20, .12	> .999
Value of sexual consent	.17	-08, .43	.332	.17	-.15, .48	.604	.004	-.30, .30	> .999
Justification of gender inequality	-.39	-.65, -.12	.001**	-.25	-.57, .07	.185	-.13	.44, -.17	.882
Rape myth acceptance	-.14	-.25, -.02	.013*	.06	-.05, .22	> .999	-.08	-.22, -.05	.433

Note. *** $p<.001$, ** $p<.01$, * $p<.05$.

(ASI) to measure baseline sexist attitudes at the start of the study. We were previously concerned about priming effects, so did not have the ASI present when retesting those who had completed it in the past in Study 1b and altogether in Study 2. However, we found consistent findings to that of Study 1a (that had the ASI present at the start) in both the null effects of exposure to sexist comments as well as the challenge rates. In this case, we could not have placed the ASI at the end of the study (after the challenge opportunity) to avoid the priming effect, since it could have been shaped by the act of challenging in this position. We therefore hypothesised:

H1: We expected that women who challenge the sexist comments will report higher self-esteem, have less support for systemic gender inequality, value sexual consent more, and accept rape myths less, in comparison to women who did not challenge.
H2: Women who scored higher in sexist attitudes will be less likely to challenge the sexist comment.
H3: Based on the argument that challenging sexism has a transformative effect, we expected that wanting to challenge sexism but not being given the possibility until the end of the study (challenge opportunity: after the study) may negatively effect outcomes for women, specifically: lower self-esteem, lower value of sexual consent, greater support for systemic gender inequality, higher rape myth acceptance, compared to those able to challenge immediately (challenge opportunity: during the study).

Method

Participants

Women were recruited via snowballing methods on social media (i.e., Facebook) as volunteers, and through the departmental student recruitment platform for course credits. 387 people began the study, and of those, 198 participants were removed for completing less than 98% of the study, and 10 were removed for not meeting the study requirements of identifying as a woman, leaving a final sample of 179 participants. Participants in this study were between the ages of 18 and 75 (M_{age} = 32, SD = 14.39). Of these women, 50% were students in a UK university, with the remaining non-students. Participants were White British (n = 102, 57%), White other (n = 41, 22.9%), other ethnicity (n = 11, 6.1%), Black British (n = 8, 4.5%), Black other (n = 7, 3.9%), Asian British (n = 6, 3.4%) and Asian other (n = 4, 2.2%). The study required a minimum of 196 participants to yield 80% power to detect a medium effect ($F2$ = 0.0625), though we had to complete the study early due to the closure of the university due to COVID-19 prevention measures.

Procedure

The design comprised a 2 (challenge opportunity: during the study versus after the study) × 2 (challenged versus did not challenge) between-subjects design. Participants were randomly allocated to the challenge opportunity (during the study or after the study conditions) at a ratio of 70:30, respectively. This ratio was due to our intention to primarily focus on the challenge immediately condition and whether challenging had a positive effect on the outcome variables (to achieve more statistical power to test a moderation effect), though we also wanted to explore the effect of challenging versus wanting to challenge.

All participants completed the study online via Qualtrics software and were advised that the goal of the study was to investigate "how you see yourself and others in society and how this impacts your performance on a cognitive task." Following informed consent, as with Studies 1a, 1b and 2, an audio clip of a male researcher introduced himself as 'John,' who provided instructions for each of the blocks in the study. 'John' then introduced the Ambivalent Sexism Inventory (ASI; Glick & Fiske, 1996), α = .85, to assess baseline sexist attitudes. An audio clip of 'John' then introduced the next block, comprising the same six items taken from the CRT-V (Sirota & Juanchich, 2018), each with four multiple choice answers. To address the potential confound of performance-related negative feedback, which may have influenced self-esteem in Studies 1a, 1b and 2, we operationalised benevolent sexism, with a new set of sexist comments embedded the CRT-V that were only encouraging, even if participants did not answer items correctly. These benevolently sexist comments were adapted from Lamarche et al. (2020), who employed language that was simultaneously patronising, infantilising, and ostensibly supportive, for example, "You seem like a very smart girl." The nature of these comments were pretested and were considered to be sexist compared to non-sexist control comments, but there was no statistically significant difference in whether participants deemed the comments sexist from answering correctly or incorrectly (see SM2 on Open Science Framework for more information: https://osf.io/2pyv4). Therefore, for each of the six items, three sexist comments were given according to whether participants answered each question correctly or incorrectly. We also included three non-sexist foil comments as a guise so that participants would feel that we were only interested in assessing their performance on the task (see Table 12 for types of feedback).

To reduce the possibility that good or poor performance was an extraneous variable that affected self-esteem, we summed participant's correct scores to include in our analyses. Participants were then randomly split whereby 70% of participants were designated to the 'challenge opportunity: during the study' condition. Firstly, participants were asked

Table 12 Sexist comment delivery for correct and incorrect answers to the CRT-V questions

Sexist Comment	Correct Answer–Sexist Comment	Incorrect Answer–Sexist Comment
Following item 1	Good answer, **clever girl**!!	This is not correct. Keep focused, you're a **clever girl** so don't let your nerves get the better of you
Following item 2	That's right! **Smart girl**!	This is not right. Don't worry, you are a **smart girl**, keep this in mind
Following item 4	You're right! **Smart girl**!	Sorry, this is not right. You are a **smart girl** but this was a tricky one
Non-Sexist Comment:	Correct Answer–Non-Sexist	Incorrect Answer-Non-Sexist
Following item 3	Smart!	Sorry, this is not right
Following item 5	Clever!	Sorry, this is not right
Following item 6	Well done! That was a tricky one	That is not correct, but that was a tricky one

Note. Text in bold is to highlight distinction between sexist and neutral comments and were not presented in bold to participants.

a filler question given to disguise the aim of the study: "*We are looking for some feedback on what you thought of the cognitive reasoning task. Overall, can you tell us what you thought?*" before being given a chance to challenge 'John' on the sexist feedback that they received by asking, "*Could you please describe your thoughts in detail on the feedback that I gave you? For instance, "Well done, smart girl!"* This provided a clear opportunity for participants to challenge the sexist comments (though we expected half of participants in this condition not to). Both questions were set to force a response, so participants could not move on until they had input a response.To examine how challenge or lack thereof influenced the outcome measures, participants then completed measures of self-esteem ($\alpha = .78$), justification for systemic gender inequality ($\alpha = .73$), attitudes towards sexual consent ($\alpha = .90$), and rape myth acceptance ($\alpha = .87$).

For the remaining 30% of the participants, they had no opportunity to challenge the sexist feedback immediately after the sexist comments and went straight to the outcome measures. We still provided an opportunity to challenge, but at the end of the study after they had completed all of the measures, so that if women chose to challenge, it would not affect their scores. Participants in this condition were unaware that they would be presented with this opportunity later. Finally, participants answered a series of sociodemographic questions before being debriefed.

Participant reactions to the comments (taken before or after the outcome variables) were independently coded by the first author and a research assistant, using the same coding schema as used in Studies 1a, 1b and 2, which indicated excellent inter-rater agreement (Cohen's $\kappa = .91$). As in Study 1a, 1b and 2, coding was then translated into superordinate codes for 'did not challenge' or 'challenged.'

Results and Discussion

The Effect of Challenging Sexist Comments

As shown in Table 13, of the 109 participants who had the opportunity to challenge the comments during the study, 65% of participants challenged the sexist comments, which was considerably higher than in Study 1a, 1b, and 2, which may be attributed to the fact that in this study, we targeted women *personally* with the comments, as opposed to women in general. We also specifically asked participants to provide their thoughts on the comments, for example, "well done, smart girl" (as opposed to providing the opportunity for general comments). In the delayed opportunity condition, 83% challenged the comments. Interestingly, the sexist comments yielded a higher rate of positive reaction than our prior studies (28% in the immediate condition and 10% in the after the study condition versus 1–4% in Studies 1a, 1b and 2) Participants' positive reactions noted the supportive tone of the comments, and for example, reported: "This definitely boosts confidence," "Made me feel smart for easy cognitive tasks, even when I got one wrong I didn't think it mattered because I was told I was still smart," and "I liked getting called a smart girl even when I failed at the questions." This

Table 13 Responses to sexist comments in immediate opportunity to challenge condition versus delayed opportunity condition in Study 3

Challenge response	Challenge Immediate ($n = 109$)	Challenge Delayed ($n = 70$)
No challenge / unclear response	$n = 8$, 7%	$n = 5$, 7%
No challenge: Positive reaction	$n = 30$, 28%	$n = 7$, 10%
Challenged	$n = 71$, 65%	$n = 58$, 83%

Note. $N = 326$.

indicates that participants found the comments in Study 3 to be encouraging.

To test H1, we examined the effect of challenging versus not challenging the sexist comments on self-esteem, value of sexual consent, justification of gender inequality, and rape myth acceptance using a MANOVA. The analysis showed that challenging was linked to lower justification of gender inequality, higher value of sexual consent, and lower rape myth acceptance. There was no statistically significant effect of challenging on self-esteem (see Table 14).

Sexist Attitudes and Challenging Sexist Comments

We conducted a logistic regression to examine H2, that sexist attitudes would predict challenging the sexist comments. Across both challenge conditions (before outcome measures and after outcome measures merged together), 72% of participants challenged the sexist comments and 28% did not. Sexist attitude did indeed statistically significantly predict challenging, Wald (1, 178) = 21.02, $p < .001$, Cox & Snell R^2 =-.14. B = -1.7 (.366).

To test whether sexist attitude influenced all participants that challenged the sexist comments, whether during or at the end of the study, we added sexist attitude as a covariate in the MANOVA. When sexist attitude was added to the model, the effects on the outcome variables were no longer significant: self-esteem F (1, 105) = 1.59, $p = .210$, $\eta_p^2 = .015$, value of sexual consent F (1, 105) = 0.07, $p = .850$, $\eta_p^2 < .001$, justification of gender inequality F (1, 105) = 1.47, $p = .228$, $\eta_p^2 = .014$, and rape myth acceptance F (1, 105) = 1.79, $p = .185$, $\eta_p^2 = 017$.

Finally, we wanted to further assess whether challenging sexism could have a transformative effect for women, or whether those that challenge are already more invested in gender equality. We therefore wanted to look at 'challengers' who challenged before answering the outcome variables and challengers who only challenged after they had already completed all measures, because we denied them the opportunity straight after the sexist comments. We conducted a MANOVA examining the effect of challenge timing (during versus after the study) on all four outcome variables. The analysis showed that being able to challenge straight away during the study was not linked with any of the dependent variables. There was a null effect of challenge timing, which supports the interpretation that challenging sexism did not itself cause shifts in participants' value of themselves and other women. Rather, it suggests that those who chose to confront the sexist comment, regardless of timing, already placed great value on women. In other words, it is not what people did that changed what they valued, but what they valued that changed how they behaved, driven by pre-existing egalitarian values (see Table 15).

General Discussion

Research on sexism suggests that sexism manifests in negative outcomes for women, but that challenging sexist comments could mitigate these outcomes, leading to calls for more research on the effects of challenging (Becker et al., 2014; Connor et al., 2017; Drury & Kaiser, 2014). The current research provides a foundation for understanding the

Table 14 Effect of challenge decision on personal and interpersonal outcomes

Outcome	Challenged	Did Not Challenge	F (1, 107)	p	η_p^2
Self-esteem	2.33 (0.43)	2.28 (0.41)	0.369	.548	.003
Justification of gender inequality	3.90 (1.23)	4.84 (1.10)	15.54***	< .001	.127
Value of sexual consent	5.86 (0.94)	5.41 (1.02)	5.48*	.021	.049
Rape myth acceptance	1.36 (0.46)	1.56 (0.52)	12.28***	< .001	.103

Note. *** $p < .001$, ** $p < .01$, * $p < .05$.

Table 15 Mean (*SD*) of the personal and interpersonal outcomes for women as a function of challenge timing (immediate opportunity to challenge/delayed opportunity to challenge) to examine the effect of 'challenging'

Outcome	Challenged Overall (N = 129)	Challenged Immediately (n = 71)	Challenged Later (n = 58)	F (1, 127)	p	η_p^2
Personal						
Self-esteem	2.34 (0.44)	2.33 (0.44)	2.35 (0.46)	0.58	.810	< .001
Value of sexual consent	5.85 (.93)	5.86 (0.94)	5.83 (0.92)	0.45	.832	< .001
Interpersonal						
Justification of gender inequality	4.09 (1.25)	3.90 (1.23)	4.32 (1.24)	3.69	.057	.028
Rape myth acceptance	1.27 (0.37)	1.26 (0.38)	1.28 (0.35)	0.94	.760	.001

associations between challenging sexist comments and immediate personal and interpersonal outcomes for women in how they value themselves and other women. To test whether challenging sexist comments benefitted women across a range of outcomes that encapsulate how women value themselves and other women, we ran three studies with the aim of distinguishing between three possibilities about the role of challenging sexist comments. First, that challenging has an immediate transformative effect of its own on women's self-value and the value they place on other women. Second, that challenging may have a moderating effect that reduces the effect of sexism. Finally, that challenging does not cause changes in women's self-value and the value they place on other women, rather, being a challenger may relate to specific psychological characteristics. We aimed to assess these possibilities using appropriate experimental and control conditions, well-powered and more diverse samples, controlling for baseline sexist attitudes, and avoiding the pressure of social responsibility by providing a private opportunity to challenge as opposed to a public one.

The first two studies (1a, 1b and 2) focused on examining the associations between receiving sexist comments on cognitive performance, self-esteem, value of sexual consent, justification of gender inequality, and rape myth acceptance, encapsulating women's personal and interpersonal value they place on themselves and other women in society. We found that across both studies, there was a null effect of receiving sexist comments on all personal and interpersonal outcomes for women. We note that this is consistent with the work of Swim and Hyers (1999), who also did not elicit an effect of exposure to sexism on performance and self-esteem. When considering the impacts of sexism, we refer to the work of Helwig (2022), noting their argument that prejudice is pervasive and persistent over time. This is also demonstrated in Swim et al.'s (2001) diary studies, in which women detail the negative impact of sexism on their self-esteem over time. The fact that we did not elicit an effect of sexism exposure is not demonstrative of the cumulative effects for women on the receiving end of insidious, systemic and prolonged sexism.

Next, we found that the act of challenging sexist comments was linked with lower support for gender inequality, more value placed on sexual consent, and lower rape myth acceptance, but there was no association between challenging sexism and a boost in self-esteem across all the studies. Previous work had focused on the effects of challenging sexism on self-esteem and wellbeing, and we added novel variables that went further in considering the impact of sexism on women's attitudes toward other women. Whilst Swim and Hyers (1999) did not find an effect of challenging on self-esteem, Gervais et al. (2010) did, though we note the absence of a non-sexist control condition in their work, thus, it is not possible to compare self-esteem of those who did not receive sexist comments and those that did. To employ a more robust methodology, we remedied this by ensuring that the current studies included a non-sexist control condition.

In Study 1 (1a and 1b), we also controlled for baseline sexist attitudes. When this was included as a covariate, the associations between challenging sexism and all outcomes became null. This suggests that the differences observed in women who challenged the sexist comments were predicted by their existing, more egalitarian attitudes towards women, rather than because of challenging the comments. We also found that baseline sexist attitudes were positively correlated with propensity to challenge, supporting the findings of Swim and Hyers (1999) in that women who are already committed to gender activism were most likely to call out sexism.

To further test whether the act of challenging sexism had transformative effects on the outcomes for women, or whether the case was that women who challenged already possessed higher value of themselves and other women, Study 3 once again tested the effects of challenging sexism. We gave an opportunity for women to challenge immediately after the sexist comments during the study, compared to delaying the opportunity to challenge the comments until post-study, after they had completed all outcome measures. This was because we expected that wanting to challenge but being denied the opportunity may impact the rumination processes, magnifying the impact of the comments. This showed a null effect in that there were no differences in personal and intrapersonal outcomes for women when they had an opportunity to challenge straight away vs. those who wanted to challenge but were not given the immediate opportunity to do so (but did later).

The results contrast with a range of findings showing immediate benefits of challenging. For instance, Gervais et al. (2010) highlighted immediate benefits of challenging sexism on competence, self-esteem, and empowerment. Similarly, Helwig (2022) found that participants who challenged sexism reported immediate greater empowerment though only when the challenge was deemed effective in changing the perpetrator's behaviour. Additionally, Foster (2015) found that women participating in collective confrontation of sexism via tweeting reported a decrease in negative affect and an increase in psychological well-being after just three days, compared to pre-study levels. While our findings did not support Gervais et al.'s (2010) claim that challenging sexism has immediate benefits on self-esteem, it is possible that the benefits of challenging sexism, much like the impact of experiencing it, accumulate over time, especially when considering Helwig's (2022) argument about the pervasive and persistent nature of sexism.

It is also important to note the significance of confrontation style. Foster (2013) found that individuals using indirect challenges to discrimination reported higher levels of

wellbeing compared to those who used an angered style. Over time, however, this trend reversed; indirect challenges led to a decrease in wellbeing, while the angered style led to an increase in wellbeing after 28 days. Longitudinal studies could track whether repeated acts of challenging sexism result in cumulative changes in self-esteem, support for gender equality, rape myth acceptance, and other personal and interpersonal outcomes, potentially uncovering effects that short-term experimental designs may miss.

Limitations

We considered that having the sexist attitude measure (ASI) at the very start of Study 1a and Study 3 may have signposted what our goal was when we later delivered the sexist comments as part of a reasoning task, and hence possibly reduced the effect of the sexist comments. To mitigate this, we retested participants in the non-sexist condition of Study 1a in Study 1b and omitted the measure altogether in Study 2 to focus on comments appearing 'out of the blue' to elicit a more natural response. We found similar challenge rates and outcomes across Studies 1a, b and 2, indicating no difference in responses when we removed the ASI from the start of the study. It is also noted that the sexist comments were more benevolent than hostile in nature, and were seen as encouraging by some women, as noted via some of the feedback. Therefore, the fact that there was no effect of exposure to these comments on personal and interpersonal outcomes may have been because the comments were not perceived as sexist "enough." However, this is unlikely, as we pretested the sexist comments and, in all cases, comments were statistically significantly rated as undermining, patronising and sexist, more so compared to non-sexist versions of the comments. We expect that, as Helwig (2022) suggests, sexism is more pervasive and persistent over time and it is very possible that as sexism is part of everyday experiences (as demonstrated in diary studies such as Swim et al. (2001) and Hyers (2007), sexist comments are more likely to have an insidious and longer-term impact over time, and there are indications that cumulative challenging over time could have personal benefits (Foster, 2013). Further, we note that in all three experiments, women that challenged could not see the benefits or costs of their challenge. Helwig (2022) found that challenge benefits to psychological wellbeing are dependent on whether the challenge was deemed effective or ineffective.

We also acknowledge the modest internal consistency of the Gender-Specific System Justification Scale (Jost & Kay, 2005; adapted from Jost & Banaji, 1994) in Study 1a, $\alpha = 0.71$, which was at the lower threshold of acceptable reliability. While this level of reliability is commonly considered adequate for research purposes, it may limit the precision of measurement and reduce the sensitivity to detect subtle effects. Future research would benefit from employing measures with higher reliability or refining existing scales to improve consistency, ensuring they are more robust.

Finally, we acknowledge that no formal attention check items were included in our studies. This decision was informed by our use of the Prolific platform for Studies 1a and 1b. Prior research demonstrates that Prolific participants provide reliable responses, even in the absence of embedded attention checks (Peer et al., 2022). The nature of participants' reactions to the sexist (and non-sexist) comments, indicates that they read and paid attention to the text. Nevertheless, we recognise the importance of incorporating formal attention checks in future studies, particularly when using alternative data collection methods or platforms, to further ensure data integrity.

Future Research Directions

Our evidence for the lack of direct immediate effect of exposure to sexist comments and of challenging those comments also require further replication and development. Several key areas warrant further investigation to refine our understanding of the potential benefits of challenging sexism. First, as our studies are limited to focusing on the *immediate* outcomes of challenging sexism, future research should examine the longer-term, cumulative effects of both experiencing and challenging sexism. Further, our studies used benevolent sexist comments. In our pretest, we confirmed that these comments were indeed perceived as sexist, though it is important to consider coping strategies to reduce dissonance when women fail to challenge these comments, where they may play down their offence to the comments to justify not speaking out (Rasinski et al., 2013). Challenging such comments may not yield the same benefits as challenging more overt forms of sexism. Future work could investigate whether challenging hostile forms of sexism leads to different outcomes.

Previous research has explored the effects of challenging sexism on self-esteem (Gervais et al., 2010; Helwig, 2022; Swim & Hyers, 1999), though self-esteem is generally considered a stable trait (Orth, 2017). To explore the broader impact of challenging, beyond self-esteem, on personal and interpersonal outcomes for women, it would be valuable to use a measure that captures both aspects, such as the Internalised Sexism Scale (Bozkur, 2020). This scale includes five key factors: self-objectification, derogation, loss of self/internalised powerlessness, competition/self-separation, and male prioritisation. Furthermore, given our findings that baseline sexist attitudes predicted challenging behaviour and that individuals with lower sexist attitudes placed higher value on themselves and other women, future research should focus on how to reduce these pervasive and enduring attitudes.

Finally, future research should investigate potential interventions aimed at amplifying the benefits of challenging sexist comments. Given our findings that challenging sexism did not have immediate transformational personal and interpersonal benefits for women, but there is an emerging evidence base supporting the value of alternative methodologies like online activism and long-term diary studies, targeted interventions could enhance the positive outcomes of challenging sexism (Foster, 2015; Hyers, 2007). Such interventions could equip women with effective strategies for confronting and processing sexist encounters, potentially fostering psychological well-being. Empowerment-based programmes or cognitive interventions that help women navigate these experiences might offer significant benefits. Furthermore, our research only focused on positive outcomes. It is possible that challenging would also have negative consequences for women, such as heighten anxiety or fears (e.g., of retaliation). In sum, while our findings provide a step forward in understanding the immediate (lack of) positive impact of challenging sexism, future research is needed to explore the long-term, contextual, and individual factors that may shape the outcomes of these experiences.

Practice Implications

These findings have practical relevance for those working to combat sexism. While challenge sexist comments did not result in immediate personal or interpersonal benefits to women, women who challenged were more likely to hold egalitarian values towards women to start with. This suggests that interventions aimed at encouraging challenging sexism may be more effective when they also focus on developing these underlying egalitarian attitudes and values. For therapists and educators, this highlights the importance of fostering positive attitudes towards women rather than relying on confrontation alone as an act of empowerment. Activists and policymakers might also reconsider the framing of challenging as immediately transformative, and efforts could focus on collectively fostering long-term gender equality and a better understanding of benevolent sexism, how to identify it, and how it serves to conceal the underlying ideological mechanisms of sexism. For everyday readers, whilst these results serve as a reminder that challenging sexism may not feel immediately transformational, it may still play a meaningful role in fostering change over time – studies have shown that challenging prejudice reduced the likelihood that perpetrators and bystanders would use negative stereotypes again in the future (Czopp et al., 2006; Rasinski & Czopp, 2010), thus reducing the prevalence of prejudice in society.

Conclusion

Our research offers new insights and a critical step forward in understanding the immediate impact of challenging sexist comments on the value women place on themselves and other women and the personal and interpersonal outcomes associated with challenging these comments. Contrary to some previous findings (e.g., Gervais et al., 2010), we observed that challenging sexism did not universally lead to immediate positive outcomes in self-esteem, value of sexual consent, justification of systemic gender inequality, and rape myth acceptance. Instead, we found that pre-existing attitudes toward sexism played a critical role whereby women with lower sexist attitudes were more likely to challenge sexist remarks and already exhibited positive characteristics in how they value themselves and other women, such as more value on sexual consent, more support for gender equality, and lower rape myth acceptance. This suggests that the act of challenging alone may not be the immediate antidote that it has previously been, but rather reflects deeper pre-existing values and beliefs.

Our findings highlight the complexity of challenging sexism and caution against overestimating its immediate benefits. While challenging sexist remarks did not produce significant short-term shifts in self-perception and perception of other women, our research shines a spotlight on the research gap in examining cumulative effects of challenging sexism over time and prompts the exploration of more long-term impacts of experiencing and challenging sexism. Overall, this work provides an important grounding for future research and practical applications in efforts to reduce sexism's pervasive impact. Continued exploration of the long-term and contextual effects of challenging, as well as interventions designed to foster cultural changes to address systemic gender inequality that underpins sexism, will be crucial in maximising the potential for meaningful change in attitudes and behaviours.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11199-025-01594-2.

Funding No funding was received to assist with the preparation of this manuscript. All authors certify that they have no affiliations with or involvement in any organisation or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

Data availability Data for all studies are available on Open Science Framework: https://doi.org/10.17605/OSF.IO/2PYV4.

References

Bareket, O., & Fiske, S. T. (2023). A systematic review of the ambivalent sexism literature: Hostile sexism protects men's power; benevolent sexism guards traditional gender roles. *Psychological Bulletin, 149*(6), 637–698. https://doi.org/10.1037/bul0000400

Baugher, S. N., Elhai, J. D., Monroe, J. R., & Gray, M. J. (2010). Rape myth acceptance, sexual trauma history, and posttraumatic stress disorder. *Journal of Interpersonal Violence, 25*(11), 2036–2053. https://doi.org/10.1177/0886260509354506

Becker, J. C. (2010). Why do women endorse hostile and benevolent sexism? The role of salient female subtypes and internalization of sexist contents. *Sex Roles, 62*(5–6), 453–467. https://doi.org/10.1007/s11199-010-9756-3

Becker, J. C., Glick, P., Ilic, M., & Bohner, G. (2011). Damned if she does, damned if she doesn't: Consequences of accepting versus confronting patronizing help for the female target and male actor. *European Journal of Social Psychology, 41*(6), 761–773. https://doi.org/10.1002/ejsp.823

Becker, J. C., Zawadzki, M. J., & Shields, S. A. (2014). Confronting and reducing sexism: A call for research on intervention. *Journal of Social Issues, 70*(4), 603–614. https://doi.org/10.1111/josi.12081

Bozkur, B. (2020). Developing internalized sexism scale for women: A validity and reliability study. *International Journal of Eurasian Education and Culture*, *5*(11), 1981–2028. https://doi.org/10.35826/ijoecc.289

Baroness Casey of Blackstock DBE CB. (2023). *Review Final Report: An independent review into the standards of behaviour and internal culture of the Metropolitan Police Service*. Retrieved January 12, 2024, from https://www.met.police.uk/SysSiteAssets/media/downloads/met/about-us/baroness-casey-review/update-march-2023/baroness-casey-review-march-2023a.pdf

Camp, J. W. (2017). *The effects of rape myth acceptance, benevolent sexism, characterological self-blame, and behavioural self-blame on posttraumatic stress disorder symptom severity* [Master's thesis, Marshall University]. Marshall Digital Scholar. https://mds.marshall.edu/etd/1108

Chaney, K. E., Young, D. M., & Sanchez, D. T. (2015). Confrontation's health outcomes and promotion of egalitarianism (C-HOPE) framework. *Translational Issues in Psychological Science, 1*(4), 363–371. https://doi.org/10.1037/tps0000042

Chaney, K. E., Sanchez, D. T., Alt, N. P., & Shih, M. J. (2020). The breadth of confrontations as a prejudice reduction strategy. *Social Psychological and Personality Science, 12*(3), 314–322. https://doi.org/10.1177/1948550620919318

Chapleau, K. M., Oswald, D. L., & Russell, B. L. (2007). How ambivalent sexism toward women and men support rape myth acceptance. *Sex Roles, 57*(1–2), 131–136. https://doi.org/10.1007/s11199-007-9196-2

Connor, R. A., Glick, P., & Fiske, S. T. (2017). Ambivalent sexism in the twenty-first century. In C. G. Sibley & F. K. Barlow (Eds.), *The Cambridge handbook of the psychology of prejudice* (pp. 295–320). Cambridge University Press.

Council of Europe. (2000). *Sexism: See it. Name it. Stop it.* Retrieved August 26, 2023, from https://human-rights-channel.coe.int/stop-sexism-en.html

Crocker, J., & Major, B. (1989). Social stigma and self-esteem: The self-protective properties of stigma. *Psychological Review, 96*(4), 608–630. https://doi.org/10.1037/0033-295X.96.4.608

Crocker, J., Voelkl, K., Testa, M., & Major, B. (1991). Social stigma: The affective consequences of attributional ambiguity. *Journal of Personality and Social Psychology, 60*(2), 218–228. https://doi.org/10.1037/0022-3514.60.2.218

Cross, E. J., & Overall, N. C. (2018). Women's attraction to benevolent sexism: Needing relationship security predicts greater attraction to men who endorse benevolent sexism. *European Journal of Social Psychology, 48*(3), 336–347. https://doi.org/10.1002/ejsp.2334

Curun, F., Taysi, E., & Orcan, F. (2017). Ambivalent sexism as a mediator for sex role orientation and gender stereotypes in romantic relationships: A study in Turkey. *Interpersona: An International Journal on Personal Relationships, 11*(1), 55–69. https://doi.org/10.5964/ijpr.v11i1.229

Czopp, A. M., Monteith, M. J., & Mark, A. Y. (2006). Standing up for a change: Reducing bias through interpersonal confrontation. *Journal of Personality and Social Psychology, 90*(5), 784–803. https://doi.org/10.1037/0022-3514.90.5.784

Dardenne, B., Dumont, M., & Bollier, T. (2007). Insidious dangers of benevolent sexism: Consequences for women's performance. *Journal of Personality and Social Psychology, 93*(5), 764–779. https://doi.org/10.1037/0022-3514.93.5.764

Dovidio, J. F., Glick, P., & Rudman, L. A. (2008). On the nature of prejudice: Fifty years after Allport. *John Wiley & Sons*. https://doi.org/10.1002/9780470773963

Drury, B. J., & Kaiser, C. R. (2014). Allies against sexism: The role of men in confronting sexism. *Journal of Social Issues, 70*(4), 637–652. https://doi.org/10.1111/josi.12083

Felmlee, D., Inara Rodis, P., & Zhang, A. (2020). Sexist slurs: Reinforcing feminine stereotypes online. *Sex Roles, 83*(1), 16–28. https://doi.org/10.1007/s11199-019-01095-z

Foster, M. (2013). The personal costs and benefits of responding to discrimination: Behavioral responses and well-being over time. *International Journal of Psychological Studies, 5*(3), 135–154. https://doi.org/10.5539/ijps.v5n3p135

Foster, M. D. (2015). Tweeting about sexism: The well-being benefits of a social media collective action. *British Journal of Social Psychology, 54*(4), 629–647. https://doi.org/10.1111/bjso.12101

Foster, M. D., Tassone, A., & Matheson, K. (2021). Tweeting about sexism motivates further activism: A social identity perspective. *British Journal of Social Psychology, 60*(3), 741–764. https://doi.org/10.1111/bjso.12431

Gervais, S. J., Hillard, A. L., & Vescio, T. K. (2010). Confronting sexism: The role of relationship orientation and gender. *Sex Roles, 63(7–8),* 463–474. https://doi.org/10.1007/s11199-010-9838-7

Glick, P., Fiske, S. T., Mladinic, A., Saiz, J. L., Abrams, D., Masser, B., Adetoun, B., Osagie, J. E., Akande, A., Alao, A., Annetje, B., Willemsen, T. M., Chipeta, K., Dardenne, B., Dijksterhuis, A., Wigboldus, D., Eckes, T., Six-Materna, I., Expósito, F., . . . López, W. L. (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

Glick, P., & Fiske, S. T. (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

Hackett, R. A., Steptoe, A., & Jackson, S. E. (2019). Sex discrimination and mental health in women: A prospective analysis. *Health Psychology, 38*(11), 1014–1024. https://doi.org/10.1037/hea0000796

Hafer, C. L., & Rubel, A. N. (2015). The why and how of defending belief in a just world. In J. M. Olson, & M. P. Zanna (Eds.), *Advances in experimental social psychology* (Vol. 51, pp. 41–96). Academic Press. https://doi.org/10.1016/bs.aesp.2014.09.001

Hammond, M. D., Overall, N. C., & Sibley, C. G. (2024). Does powerlessness motivate men and women to endorse sexism? *European Journal of Social Psychology, 54*(6), 1354–1363. https://doi.org/10.1002/ejsp.3092

Heilman, M. E. (2001). Description and prescription: How gender stereotypes prevent women's ascent up the organizational ladder. *Journal of Social Issues, 57*(4), 657–674. https://doi.org/10.1111/0022-4537.00234

Helwig, S. (2022). *The role of confrontation effectiveness for women's psychological wellbeing after a sexist event* [Doctoral dissertation, The University of Maine]. DigitalCommons@UMaine. https://digitalcommons.library.umaine.edu/etd/3668

Herrero, J., Rodríguez, F. J., & Torres, A. (2017). Acceptability of partner violence in 51 societies: The role of sexism and attitudes toward violence in social relationships. *Violence against Women, 23*(3), 351–367. https://doi.org/10.1177/1077801216642870

Hershcovis, M. S., Cameron, A.-F., Gervais, L., & Bozeman, J. (2018). The effects of confrontation and avoidance coping in response to workplace incivility. *Journal of Occupational Health Psychology, 23*(2), 163–174. https://doi.org/10.1037/ocp0000078

Hopkins-Doyle, A., Sutton, R. M., Douglas, K. M., & Calogero, R. M. (2019). Flattering to deceive: Why people misunderstand benevolent sexism. *Journal of Personality and Social Psychology, 116*(2), 167–192. https://doi.org/10.1037/pspa0000135

Humphreys, T. P., & Brousseau, M. M. (2010). The Sexual Consent Scale–Revised: Development, reliability, and preliminary validity. *Journal of Sex Research, 47*(5), 420–428. https://doi.org/10.1080/00224490903151358

Hyers, L. L. (2007). Resisting prejudice every day: Exploring women's assertive responses to anti-Black racism, anti-Semitism, heterosexism, and sexism. *Sex Roles, 56*(1–2), 1–12. https://doi.org/10.1007/s11199-006-9142-8

Jones, K., Stewart, K., King, E., Botsford Morgan, W., Gilrane, V., & Hylton, K. (2014). Negative consequence of benevolent sexism on efficacy and performance. *Gender in Management: An International Journal, 29*(3), 171–189. https://doi.org/10.1108/gm-07-2013-0086

Jost, J. T., & Banaji, M. R. (1994). The role of stereotyping in system-justification and the production of false consciousness. *British Journal of Social Psychology, 33*(1), 1–27. https://doi.org/10.1111/j.2044-8309.1994.tb01008.x

Jost, J. T., & Kay, A. C. (2005). Exposure to benevolent sexism and complementary gender stereotypes: Consequences for specific and diffuse forms of system justification. *Journal of Personality and Social Psychology, 88*(3), 498–509. https://doi.org/10.1037/0022-3514.88.3.498

Lamarche, V. M., Seery, M. D., Kondrak, C. L., Saltsman, T. L., & Streamer, L. (2020). Clever girl: Benevolent sexism and cardiovascular threat. *Biological Psychology, 149,* Article 107781. https://doi.org/10.1016/j.biopsycho.2019.107781

Landrine, H., & Klonoff, E. A. (1997). *Discrimination against women: Prevalence, consequences, remedies* (ED417275). ERIC. https://eric.ed.gov/?id=ED417275

Lerner, M. J. (1980). The belief in a just world. In *The belief in a just world: A fundamental delusion* (pp. 9–30). Springer. https://doi.org/10.1007/978-1-4899-0448-5_2

Lonsway, K. A., & Fitzgerald, L. F. (1994). Rape myths: In review. *Psychology of Women Quarterly, 18*(2), 133–164. https://doi.org/10.1111/j.1471-6402.1994.tb00448.x

Lorenz, K., Kirkner, A., & Ullman, S. E. (2019). A qualitative study of sexual assault survivors' post-assault legal system experiences. *Journal of Trauma & Dissociation, 20*(3), 263–287. https://doi.org/10.1080/15299732.2019.1592643

Major, B., Gramzow, R. H., McCoy, S. K., Levin, S., Schmader, T., & Sidanius, J. (2002). Perceiving personal discrimination: The role of group status and legitimizing ideology. *Journal of Personality and Social Psychology, 82*(3), 269–282. https://doi.org/10.1037/0022-3514.82.3.269

Mallett, R. K., & Wagner, D. E. (2011). The unexpectedly positive consequences of confronting sexism. *Journal of Experimental Social Psychology, 47*(1), 215–220. https://doi.org/10.1016/j.jesp.2010.10.001

National Education Union & UK Feminista. (2017). *"It's just everywhere" A study on sexism in schools and how we tackle it*. Retrieved March 23, 2023, from https://ukfeminista.org.uk/wp-content/uploads/2017/12/Report-Its-just-everywhere.pdf

Noh, S., & Kaspar, V. (2003). Perceived discrimination and depression: Moderating effects of coping, acculturation, and ethnic support. *American Journal of Public Health, 93*(2), 232–238. https://doi.org/10.2105/AJPH.93.2.232

Office for National Statistics. (2020). *Nature of sexual assault by rape or penetration, England and Wales: Year ending March 2020*. Retrieved August 28, 2023 from https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/articles/natureofsexualassaultbyrapeorpenetrationenglandandwales/yearendingmarch2020

Orth, U. (2017). The lifespan development of self-esteem. In J. Specht (Ed.), *Personality development across the lifespan* (pp. 181–195). Elsevier Academic Press. https://doi.org/10.1016/B978-0-12-804674-6.00012-0

Ortiz, S. M. (2023). "If Something Ever Happened, I'd Have No One to Tell": How online sexism perpetuates young women's silence. *Feminist Media Studies, 24*(1), 119–134. https://doi.org/10.1080/14680777.2023.2185565

UK Parliament. (2024). *Sexism in the city - written evidence*. Retrieved May 13, 2024, from https://committees.parliament.uk/work/7842/sexism-in-the-city/publications/written-evidence/

Payne, D. L., Lonsway, K. A., & Fitzgerald, L. F. (1999). Rape myth acceptance: Exploration of its structure and its measurement using the Illinois rape myth acceptance scale. *Journal of Research in Personality, 33*(1), 27–68. https://doi.org/10.1006/jrpe.1998.2238

Peer, E., Rothschild, D., Gordon, A., Evernden, Z., & Damer, E. (2022). Data quality of platforms and panels for online behavioral research. *Behavior Research Methods, 54*(4), 1643–1662. https://doi.org/10.3758/s13428-021-01694-3

Peterson, Z. D., & Muehlenhard, C. L. (2004). Was it rape? The function of women's rape myth acceptance and definitions of sex in labeling their own experiences. *Sex Roles, 51*(3), 129–144. https://doi.org/10.1023/B:SERS.0000037758.95376.00

Plan International. (2017). *Young people more likely to be feminist.* Retrieved March 14, 2024, from https://plan-uk.org/media-centre/young-people-more-likely-to-be-feminist-research#:~:text=Girls'%20rights%20charity%20Plan%20International,more%20likely%20to%20do%20so

Rasinski, H. M., & Czopp, A. M. (2010). The effect of target status on witnesses' reactions to confrontations of bias. *Basic and Applied Social Psychology, 32*(1), 8–16. https://doi.org/10.1080/01973530903539754

Rasinski, H. M., Geers, A. L., & Czopp, A. M. (2013). "I guess what he said wasn't that bad" dissonance in nonconfronting targets of prejudice. *Personality and Social Psychology Bulletin, 39*(7), 856–869. https://doi.org/10.1177/0146167213484769

Riverlight. (2024). *In the judge's words*. Retrieved June 20, 2024, from https://www.riverlight.org.uk/campaign?ss_source=sscampaigns&ss_campaign_id=6640d11c97f32531270b4854&ss_email_id=6640da9391c2ca04673c1068&ss_campaign_name=Riverlight+

Newsletter+-+May+2024&ss_campaign_sent_date=2024-05-12T15%3A05%3A00Z

Romero-Sánchez, M., Durán, M., Carretero-Dios, H., Megías, J. L., & Moya, M. (2010). Exposure to sexist humor and rape proclivity: The moderator effect of aversiveness ratings. *Journal of Interpersonal Violence, 25*(12), 2339–2350. https://doi.org/10.1177/088626050935488

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton University Press.

Sanchez, D. T., Himmelstein, M. S., Young, D. M., Albuja, A. F., & Garcia, J. A. (2015). Confronting as autonomy promotion: Speaking up against discrimination and psychological well-being in racial minorities. *Journal of Health Psychology, 21*(9), 1999–2007. https://doi.org/10.1177/1359105315569619

Sibley, C. G., Overall, N. C., & Duckitt, J. (2007). When women become more hostilely sexist toward their gender: The system-justifying effect of benevolent sexism. *Sex Roles, 57*, 743–754. https://doi.org/10.1007/s11199-007-9306-1

Sirota, M., & Juanchich, M. (2018). Effect of response format on cognitive reflection: Validating a two-and four-option multiple choice question version of the Cognitive Reflection Test. *Behavior Research Methods, 50*, 2511–2522. https://doi.org/10.3758/s13428-018-1029-4

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35*(1), 4–28. https://doi.org/10.1006/jesp.1998.1373

Swim, J. K., & Hyers, L. L. (1999). Excuse me—What did you just say?!: Women's public and private responses to sexist remarks. *Journal of Experimental Social Psychology, 35*(1), 68–88. https://doi.org/10.1006/jesp.1998.1370

Swim, J. K., Hyers, L. L., Cohen, L. L., & Ferguson, M. J. (2001). Everyday sexism: Evidence for its incidence, nature, and psychological impact from three daily diary studies. *Journal of Social Issues, 57*(1), 31–53. https://doi.org/10.1111/0022-4537.00200

The Fawcett Society. (2018). *Legal system 'failing' women and in need of reform, says Fawcett in landmark sex discrimination law review*. Retrieved October 15, 2023, from https://www.fawcettsociety.org.uk/news/legal-system-failing-women-need-reform-says-fawcett-landmark-sex-discrimination-law-review

Tiggemann, M., & McGill, B. (2004). The role of social comparison in the effect of magazine advertisements on women's mood and body dissatisfaction. *Journal of Social and Clinical Psychology, 23*(1), 23–44. https://doi.org/10.1521/jscp.23.1.23.26991

Topping, A. (2021). Finally the UK has noticed its rampant sexism in healthcare. What now? *The Guardian.* Retrieved August 4, 2024, from https://www.theguardian.com/society/2021/dec/23/finally-the-uk-has-noticed-its-rampant-sexism-in-healthcare-what-now

Turner, A. (2024). Sexism and misogyny as traits of police culture: Problems, red flags and solutions. *International Journal of Police Science & Management, 26*(2), 279–291. https://doi.org/10.1177/14613557241228736

United Nations. (2024). *Sustainable development goals: Goal 5: Achieve gender equality and empower all women and girls.* Retrieved May 20, 2024, from https://www.un.org/sustainabledevelopment/gender-equality/

Women and Equalities Committee. (2024). *Misogyny in music.* UK Parliament. Retrieved May 17, 2024, from https://publications.parliament.uk/pa/cm5804/cmselect/cmwomeq/129/report.html

Woodzicka, J. A., & LaFrance, M. (2001). Real versus imagined gender harassment. *Journal of Social Issues, 57*(1), 15–30. https://doi.org/10.1111/0022-4537.00199

Zawisza, M., Kosakowska-Berezecka, N., Glick, P., Olech, M., Besta, T., Jurek, P., Sobiecki, J., Best, D. L., Bosson, J. K., Vandello, J. A., Safdar, S., Włodarczyk, A., & Żadkowska, M. (2025). *Worse for women, bad for all: A 62-nation study confirms and extends ambivalent sexism principles to reveal greater social dysfunction in sexist nations*. Advance online publication. https://doi.org/10.1177/19485506241302882