JOURNAL OF E-LEARNING AND KNOWLEDGE SOCIETY
Vol. 22, No. 1 (2026)

# When AI is sexist: Investigating responses to AI and human discrimination

Selen Akay[a,1], Junko Kanero[a], Sabahat Cigdem Bagci[a]

[a]*Faculty of Arts and Social Sciences, Sabancı University – Istanbul (Türkiye)*

*(submitted: 10/03/2026; accepted: 9/6/2026; published: 23/7/2026)*

**Abstract**

As artificial intelligence (AI) systems prove their utility in various aspects of daily life, their potential risk to societal progress is still debated given the increasing rates of their discriminatory behaviors. Some preliminary evidence suggests that the same discriminatory behavior may be perceived differently when performed by a person or an AI system, yet it remains largely unknown how people perceive and respond to AI's discriminatory behaviors. Through two studies conducted in Türkiye (Study 1) and the United States (Study 2), participants read about either AI's discrimination against women, humans' discrimination against women, or an unrelated topic, and indicated the degree of perceived threat to women, willingness to participate in collective action for women's equal treatment, and their attitudes toward women and feminists. In both cultures, discrimination was perceived to be less threatening to women when committed by AI compared to a human, yet collective action intentions and attitudes toward women and feminists did not differ across conditions. Notably, American men consistently reported less favorable attitudes toward women after reading about discrimination. These findings suggest discrimination by AI requires special attention and underline the importance of investigating the influence of technologies on sustaining intergroup inequality.

**KEYWORDS:** discrimination, intergroup threat, artificial intelligence (AI), collective action

**DOI**
https://doi.org/10.20368/1971-8829/1136320

**CITE AS**
Akay, S., Kanero, J., & Bagci,S.C. (2026). When AI is sexist: Investigating responses to AI and human discrimination. *Journal of e-Learning and Knowledge Society*, *22*(1). https://doi.org/10.20368/1971-8829/1136320

## 1. Introduction

Recent technological advancements have led artificial intelligence (AI) tools to be available for the use of the general public. AI tools can perform a wide range of tasks, from writing elaborate essays to create workout routines and giving relationship advice. Considering the availability and popularity of AI technologies, the ethical implications of interacting with AI systems, as well as the potential risks of incorporating them into our lives, have become a central topic of debate (e.g., Darling, 2012, 2015). While AI technologies have the potential to reduce social inequality by providing affordable tools to perform tasks that are otherwise challenging for disadvantaged groups, they can also exacerbate inequality in various ways, particularly by discriminating against already disadvantaged populations.

Technologies often make seemingly avoidable mistakes or errors, ranging from technical errors to social norm violations (Giuliani et al., 2015; Tian & Oviatt, 2021). Discrimination – the inappropriate and unfair negative treatment of others based on their group identities (Allport, 1954; Dovidio et al., 2010) – is one of such behaviors that AI systems occasionally display based on the user's characteristics or the input they give to the AI. One explanation for AI discrimination is algorithmic bias, which arises when algorithms are trained on data that reflect existing social biases (Danks & London, 2017). Such biases may be encoded intentionally by developers or emerge unintentionally from exclusionary datasets that overrepresent Western, White, and male populations (Bender et al., 2021; Hall & Ellis, 2023). Because algorithms tend to overgeneralize patterns in available data, they can produce biased inferences – such as underestimating women's suitability for STEM roles or using postcodes as proxies for race in crime-risk assessments – thereby reinforcing existing inequalities even when datasets appear neutral (Dastin, 2018; Allen & Masters, 2020)

[1] corresponding author - email: selen.akay@sabanciuniv.edu – address: Orta Mahalle, Universite Caddesi No:27, 34956 Tuzla, Istanbul, Turkiye

and disproportionately impacting disadvantaged groups such as women, racial/ethnic minorities, and sexual minorities (e.g., Ferrara, 2023). While perceiving discrimination can adversely affect targets by reducing their well-being (e.g., Bagci et al., 2020), and damaging their physical and mental health (e.g., Borrell et al., 2007; Landrine & Klonoff, 1996; Pascoe & Smart Richman, 2009; Schmitt et al., 2014; Williams et al., 2003), exposure to AI discrimination may also shape both disadvantaged and advantaged group members' willingness to take action against such unfair treatment. The current studies investigate this assumption and test the implications of AI discrimination (compared to human discrimination) for advantaged and disadvantaged group members' evaluation of the target group and the extent at which they are willing to fight for disadvantaged groups' rights.

A first step to understand reactions to AI discrimination is to understand how group members respond to discrimination experiences across status, irrespective of the source. One robust finding is that advantaged and disadvantaged groups perceive discrimination and group-based inequality differently (e.g., Knowles & Lowery, 2012). Due to power imbalances and structural limitations, disadvantaged or low-status groups are more frequently stigmatized and discriminated against (e.g., Link & Phelan, 2001) and the perception of discrimination is often a more internal, uncontrollable, and chronic social experience (Schmitt et al., 2002). In contrast, advantaged groups are less likely to perceive discrimination compared to disadvantaged groups, and they are less likely to challenge the status quo (e.g., Brown et al., 2022; Caricati & Owuamalam, 2020; Lowery et al., 2007; Sidanius & Pratto, 1999), and may even engage in collective action that aim to maintain the existing system favoring group-based inequality (e.g., Guvensoy et al., 2025; Osborne et al., 2019).

Among disadvantaged group members, seeking social change through collective action is one of the most effective responses to inequality (Ellemers et al., 1990; Wright et al., 1990). Collective action research suggests that there are many drivers of collective action behaviors, but one of the main motivations is perceived injustice and discrimination, often fueling the emotion of anger among disadvantaged group members (e.g., Van Zomeren et al., 2008). Indeed, research shows that people's recognition of group-based disadvantages drives people to challenge discrimination (Ellemers & Barreto, 2009; Wright et al., 1990; Wright & Tropp, 2002). For example, Uluğ and colleagues (2023) demonstrated that women who have witnessed gender discrimination through an experimental procedure were more willing to engage in collective action for gender equality. Thus, exposure to discrimination targeting one's own disadvantaged group is likely to increase disadvantaged group members' collective action intentions.

Moreover, both theoretical and empirical work suggests that perceiving discrimination or similar experiences of devaluation or exclusion targeting the ingroup may lead disadvantaged group members to identify with their ingroup more, thereby protecting themselves from the adverse effects of discrimination (Branscombe et al., 1999). Findings suggest that identifying with one's ingroup after perceiving group-based prejudice and discrimination may lead disadvantaged groups to evaluate their ingroup more positively (Dion & Earn, 1975). Thus, exposure to AI's discrimination against disadvantaged groups may potentially improve disadvantaged group members' attitudes toward the targeted group and their advocates (e.g., political groups such as feminists) and increase their willingness to engage in collective action for their equal treatment.

In contrast to disadvantaged groups, advantaged group members may either support equality struggles as allies or resist them to preserve the status quo (Kutlaca et al., 2022; Flood et al., 2021). Collective action research has traditionally portrayed advantaged groups as less supportive of equality, largely because they are motivated to maintain their privileges and perceive social change as a zero-sum loss (e.g., Sidanius & Pratto, 1999; Caricati & Owuamalam, 2020; Brown et al., 2022). Accordingly, advantaged groups are less likely to recognize discrimination against disadvantaged groups and may deny inequality when confronted with their own privilege (Knowles & Lowery, 2012; Phillips & Lowery, 2015), a pattern illustrated by findings showing that men and White individuals are less likely to view gender and racial inequalities as unjust (Miron et al., 2011).

Nevertheless, advantaged group members are also likely to respond to discrimination targeting disadvantaged group members in various ways. For example, research demonstrates that advantaged groups who perceive other groups as relatively deprived are more likely to support pro-equality policies and affirmative actions that help the disadvantaged (e.g., Mazzuca et al., 2022; Tougas & Veilleux, 1990; Veilleux & Tougas, 1989). Previous research has shown evidence that exposure to disadvantaged group members' discrimination may lead advantaged group members to display more positive attitudes toward the discriminated outgroup (Bagci et al., 2017) and greater motivation to engage in collective action supporting intergroup equality (Uluğ & Tropp, 2021), in the context of ethnic and racial relationships. Notably, some findings indicate that inequality, historical mistreatment, and conflict with disadvantaged groups can lead advantaged groups to feel more positively about the disadvantaged and reprimand their mistreatment, due to feelings such as guilt and shame (e.g., Allpress et al., 2010, 2014; Brown & Cehajic, 2008; Swim & Miller, 1999). Thus, it could be expected that advantaged groups would feel more positively about disadvantaged groups and their advocates, and would be more likely to engage in collective action supporting intergroup equality after being exposed to the target group's discrimination.

Overall, perceiving discrimination and group-based inequalities may make both advantaged and disadvantaged groups more willing to engage in collective action for equality and improve their attitudes toward the disadvantaged and their advocates. Nevertheless, previous research has barely touched upon whether and how discrimination by AI would be evaluated by both advantaged and disadvantaged group members. Although responses to AI discrimination may parallel familiar intergroup processes, AI and related technologies are treated as novel social entities (Gray et al., 2025; Kahn & Shen, 2017), highlighting the importance of understanding the unique dynamics and consequences of this form of discrimination.

As AI systems become incorporated into domains that influence people's lives, such as finance, education, and job recruitment, concerns about AI discrimination have become more relevant than ever. However, since AI discrimination is only a recent concern, it is unclear whether it triggers the same reactions as human discrimination. One reason why AI's discriminatory behaviors may affect intergroup relationships differently than humans' discriminatory behaviors is because AI and similar technologies are not attributed as much mind (i.e., agentic and experiential capabilities) as humans (Gray et al., 2007). Therefore, one potential scenario is that people may perceive AI as less of a threat because, while AI and similar technologies are attributed some moral responsibility and intentionality, they are perceived to be less agentic compared to humans (Bigman et al., 2019; Floridi & Sanders, 2004; Geiselmann et al., 2023; Shank & DeSanti, 2018; Shank et al., 2021), overall making their discrimination relatively less intentional and thereby less threatening compared to human discrimination. This assumption also suggests that perceiving AI discrimination as less threatening may be also less likely to motivate both advantaged and disadvantaged group members' support and willingness to fight on behalf of disadvantaged group members' rights.

Supporting this suggestion, empirical evidence demonstrates that AI's discriminatory behaviors and moral violations are evaluated differently from humans' discriminatory behaviors and moral violations (Bigman et al., 2023; Shank & DeSanti, 2018). For instance, Bigman and colleagues (2023) compared the extent to which participants felt moral outrage (i.e., anger or disgust triggered by perceived violations of moral values) due to the discriminatory behaviors of algorithms and humans in a hiring setting. Across both racial and gender contexts, algorithms' discriminatory behaviors evoked less moral outrage in participants, and participants attributed less prejudice to the algorithm than they did to the human recruiter. Thus, even when a person and an AI system engage in the same discriminatory behavior, resulting in the same negative consequences, the behavior of AI discriminators are likely to be dismissed. However, the dismissal of AI discriminators may not necessarily indicate that their discriminatory behaviors are inconsequential since such biases can lead to the perpetuation of harmful stereotypes (e.g., Vlasceanu & Amodio, 2022) and make discrimination of particular disadvantaged groups more normative and legitimate.

Overall, AI's biases and discriminatory tendencies may lead to awkward, tense, or outrageous situations, perpetuating harmful stereotypes. While some work compared people's reactions to discrimination by AI's and humans (Bigman et al., 2023), no research investigated whether AI's discriminatory behaviors against disadvantaged groups affect people's attitudes toward the target group and their motivation to engage in collective action.

The current research uses gender inequality in two different countries as a context to investigate whether reading about AI's discriminatory behaviors against disadvantaged groups (i.e., women) influences people's collective action intentions for the disadvantaged group's equality, and their attitudes toward feminists (Study 1) and women (Study 2). Gender inequality is a particularly relevant context for examining AI discrimination because of its universality and commonality (United Nations Development Programme, 2023) and the marked gender imbalance among AI developers (Wajcman, 2010; Vargas-Solar, 2022).

Although gender discrimination is a global issue, reactions to gender inequality may vary across cultural contexts. Türkiye (Study 1) ranks relatively low on global gender equality indices, and gender discrimination against women is widely documented as structurally embedded in society (Kandiyoti, 2016; Özdemir Sarıgil, 2022; UNDP, 2023; World Economic Forum, 2024). At the same time, AI systems are increasingly integrated into sectors such as education and healthcare, and despite slower overall adoption compared to other countries, the Turkish population shows relatively high levels of AI use and familiarity (Eurostat, 2025; Maslej et al., 2025). Given Türkiye's persistent gender inequality alongside growing exposure to AI, it offers a relevant context for examining reactions to AI-driven gender bias. On the other hand, the United States (Study 2) ranks much higher in gender equality than Türkiye (World Economic Forum, 2025) and is home to several major AI research and development companies, including OpenAI, Google, and Meta, and AI technologies are widely deployed across multiple sectors (e.g., McElheran et al., 2024). Nevertheless, many documented cases of AI discrimination involve US-based companies or residents (e.g., Dastin, 2018; U.S. Equal Employment Opportunity Commission, 2022), providing two unique socio-cultural contexts for the investigation of the current research questions.

Based on the perceived discrimination literature, we expected that, irrespective of the source, perceived discrimination is likely to improve attitudes toward the target group and motivate both advantaged and disadvantaged group members to support women's rights. Considering that both advantaged (Batson et al., 1997, 2002; Swim & Miller, 1999; Vescio et al., 2003) and disadvantaged groups (Becker & Barreto, 2014; Dion & Earn, 1975; Osborne & Sibley, 2015; Uluğ et al., 2023) are likely to approve egalitarian policies and feel more positively about the targeted group and targets who confront prejudice, they may also feel positively about people who advocate for the disadvantaged (i.e., feminists). Therefore, we expected that:

- H1: Compared to participants in the control condition, participants in the human and AI discrimination conditions would both perceive greater threat to women, show more positive attitudes toward feminists (Study 1) and women (Study 2), and greater motivation to engage in collective action supporting women's rights.

Based on findings that demonstrate AI and similar technologies to be attributed less capacity for intentional behavior and prejudice as humans (Bigman et al., 2023; Gray et al., 2007; Shank & DeSanti, 2018), and since AI's discrimination evokes less moral outrage than humans' discrimination (Bigman et al., 2023), we expected that people would perceive AI discrimination as less threatening (harmful) to women than human discrimination:

- H2: Compared to the AI discrimination condition, participants in the human discrimination condition would perceive greater threat to women, show more positive attitudes toward feminists (Study 1) and women (Study 2), and greater motivation to engage in collective action supporting women's rights.

## 2. Study 1

Study 1 was conducted in Türkiye and examined both advantaged (men) and disadvantaged (women) group members' reactions to AI discrimination, compared to human discrimination and the control condition.

### 2.1 Materials and Method

#### *2.1.1 Participants and Procedure*

We recruited 350 participants through both convenience sampling and a research participation system at Sabanci University. A priori power analysis with G*Power (Faul et al., 2007) suggested at least 330 participants to reach 80% power with $\alpha = .05$ and a small effect size (estimated based on Bigman et al., 2023).

Six participants who did not identify as men or women were removed from all analyses as this study focuses on contrasting the advantaged (men) and disadvantaged groups (women). As the duration data were heavily skewed, we first performed a log10 transformation to identify outliers and then removed participants who took 3 SDs above average. Seven participants who took more than 109 minutes to complete the study were also removed from the data. The final sample consisted of 337 native Turkish-speaking adults ($M_{age} = 33.43$, $SD = 14.91$, 233 women). Participants reported having an average income ($M = 4.77$, $SD = 1.09$, ranging from 1 = *Low* to 7 = *High*), being moderately informed about AI systems ($M = 4.14$, $SD = 1.60$), and having an average amount of experience using AI systems ($M = 4.16$, $SD = 2.05$).

Participants were assigned to one of the three conditions: AI discriminator, human discriminator, or control condition. In the AI discriminator and human discriminator conditions, participants read a fake news report about AI systems or humans discriminating against women. In the human discriminator condition, the source of the discrimination was identified as "industry experts". Participants in the control condition read a fake news report on an unrelated topic. After reading the news report, participants responded to a questionnaire consisting of a manipulation check, scales about perceived threats toward women, collective action intentions for women's equal treatment, and their attitudes toward feminists, demographic questions, and a suspicion check. Students recruited through the course credit system were awarded course credits for their participation.

#### *2.1.2 Materials*

Fake news report. Fake news reports discussed either AI systems' discriminatory behaviors toward women, industry experts' discriminatory behaviors toward women, or an unrelated text about the agriculture sector. The reports were created based on examples of discrimination available from online sources and existing news articles that discuss AI's discriminatory behaviors (Dastin, 2018; Jonker & Rogers, 2024), and a team of researchers evaluated the texts to ensure the clarity of the writing and the manipulation (see Supplementary Materials).

Manipulation check. Participants rated the extent to which they perceived discrimination against women in the news report, on a scale of 1 (*Strongly Disagree*) to 7 (*Strongly Agree*).

Perceived threat toward women. The measure consisted of three items (e.g., '[Artificial intelligence systems/industry experts/the agriculture sector] can damage women's physical, mental, or emotional health'), all rated on a scale of 1 (*Strongly Disagree*) to 7 (*Strongly Agree*). AI systems ($\alpha = .89$), industry experts ($\alpha = .91$), and the agriculture sector ($\alpha = .92$) all demonstrated good reliability.

Collective action intentions. Collective action intentions for women's equal treatment were measured with three items assessing their willingness to donate, attend a protest, and sign a petition for women's equal treatment, on a scale of 1 (*Very Unlikely)* to 7 (*Very Likely*). The scale was adapted from prior normative collective action measures used in the literature (e.g., Hoskin et al., 2019), and demonstrated good reliability ($\alpha = .73$).

Attitudes toward feminists. Attitudes toward feminists were measured with 6 bipolar items with a 7-point scale (e.g., 1 = *Contempt*, 7 = *Respect*) (e.g., Bagci et al., 2020; Wright et al., 1997), with higher scores indicating more positive attitudes ($\alpha = .92$).

Demographics. Participants indicated their age, gender, ethnicity, economic status, and knowledge about and experience with AI systems.

## 2.2 Results

### *2.2.1 Manipulation Check*

A two-way ANOVA with Condition and Gender as predictors resulted in the significant effect of Condition ($F(2, 331) = 53.05$, $p < .001$, $\eta_p^2 = .243$). Tukey's HSD tests revealed that participants in the AI discriminator condition ($M = 5.80$, $SD = 1.63$) perceived discrimination against women to be higher than those in the control condition ($M = 3.70$, $SD = 2.02$; $p < .001$), but did not differ from those in the human discriminator condition ($M = 6.06$, $SD = 1.37$; $p = .468$). Moreover, Gender significantly influenced discrimination perceptions ($F(1, 331) = 10.32$, $p = .001$, $\eta_p^2 = .030$). Specifically, female participants ($M = 5.48$, $SD = 1.92$) perceived more discrimination against women than did male participants ($M = 4.64$, $SD = 1.99$). No interaction effect was found ($p = .607$).

### *2.2.2 Main Effects of Condition*

A MANOVA with Condition as the predictor assessed whether there are significant differences across conditions regarding perceived threat, collective action intentions for women's equal treatment, and attitudes toward feminists. Multivariate tests showed that Condition had a significant effect ($F(6, 666) = 16.78$, $p < .001$, $\eta_p^2 = .131$), and the univariate tests revealed the effect of Condition on threat perceptions toward women ($F(2, 334) = 56.19$, $p < .001$, $\eta_p^2 = .252$). Tukey's HSD tests showed that perceived threat toward women was higher in the human discriminator condition ($M = 5.56$, $SD = 1.53$) compared to the AI discriminator ($M = 4.52$, $SD = 1.54$; $p < .001$) and control conditions ($M = 3.26$, $SD = 1.83$; $p < .001$). Similarly, perceived threat in the AI discriminator condition was significantly higher than control ($p < .001$). Collective action intentions for women's equal treatment ($p = .681$) and attitudes toward feminists ($p = .096$) did not significantly differ across conditions. These findings were replicated across the student sample and the community sample.

### *2.2.3 Exploratory Analyses*

An exploratory MANOVA tested whether participants' gender predicted perceived threat to women, collective action intentions, and attitudes. The multivariate effects of Condition ($F(6, 660) = 13.07$, $p < .001$, $\eta_p^2 = .106$) and Gender were significant ($F(3, 329) = 24.55$, $p < .001$, $\eta_p^2 = .183$), but their interaction was not ($p = .138$). The unique effect of Condition on perceived threat toward women remained significant ($F(2, 331) = 41.98$, $p < .001$, $\eta_p^2 = .202$). Condition did not affect collective action intentions ($p = .605$) nor attitudes toward feminists ($p = .213$).

Gender significantly and uniquely affected perceived threat toward women ($F(1, 331) = 10.75$ $p = .001$, $\eta_p^2 = .031$), collective action intentions for women's equal treatment ($F(1, 331) = 32.91$, $p < .001$, $\eta_p^2 = .090$), and attitudes toward feminists ($F(1, 331) = 57.37$, $p < .001$, $\eta_p^2 = .148$). Female participants perceived significantly more threat ($M = 4.72$, $SD = 1.87$) than male participants ($M = 3.95$, $SD = 1.82$). Similarly, female participants' collective action intentions ($M = 5.96$, $SD = 1.06$) and attitudes toward feminists ($M = 5.28$, $SD = 1.33$) were significantly higher than male participants' collective action intentions ($M = 5.14$, $SD = 1.39$) and attitudes toward feminists ($M = 4.03$, $SD = 1.45$). We found no interaction effects (*p's* > .100).

Lastly, two exploratory mediation analyses were conducted via the Process Macro (Hayes, 2022; Model 4) to see whether perceived threat mediated the effect of condition (IV) on collective action intentions for women's equal treatment and attitudes toward feminists (DV). Indeed, perceived threat mediated the effect of condition on collective action intentions for women's equal treatment ($B = .10$, $SE = .04$, 95% CI [0.02, 0.18]). Moreover, the relative indirect effect of the human discriminator condition ($B = .22$, $SE = .11$, 95% CI [0.01, 0.45]) was stronger than that of the AI discriminator condition ($B = .12$, $SE = .06$, 95% CI [0.00, 0.26]), as indicated by the coefficients. Similarly, perceived threat mediated the effect of condition on attitudes toward feminists ($B = .23$, $SE = .09$, 95% CI [0.06, 0.43]). Again, the relative indirect effect of the human discriminator condition ($B = .39$, $SE = .14$, 95% CI [0.14, 0.67]) was stronger than that of the AI discriminator condition ($B = .22$, $SE = .08$, 95% CI [0.07, 0.39]), as indicated by the coefficients.

## 2.3 Discussion

The results suggest that, while discrimination by AI and humans was perceived as a potential threat to women (partial support for H1), AI discrimination was perceived as a lower threat compared to human discrimination. This finding is in line with existing evidence indicating that AI is believed to be less agentic (Gray et al., 2007) and evokes less moral outrage and is attributed less prejudice (Bigman et al., 2023). Unlike our expectations, however, collective action intentions or attitudes toward feminists remained unaffected by the news report about discrimination. This finding was surprising, as prior research suggests that, in response to unequal treatment, both targets and non-targets become more willing to engage in collective action for disadvantaged groups (Mazzuca et al., 2022; Moscatelli et al., 2025), and report more positive attitudes toward discriminated groups (Bagci et al., 2017). While we did not find significant differences across conditions, we found that perceived threat mediated the effect of condition on collective action intentions as well as attitudes. Although the effect sizes are small, these results indicate that threat from human and AI discriminators relate to intergroup outcomes.

## 3. Study 2

Study 1 established that AI discrimination yields lower threat than human discrimination, yet the conclusion should not be drawn based on a single experiment. To evaluate the reliability and generalizability of the initial results and examine potential cross-cultural variation, Study 2 tested participants in the US. As discussed in the introduction, Türkiye ranks relatively low in global indices of gender equality and AI-related development. One possibility is that this context contributed to AI being perceived as less threatening to women in Study 1. The US, on the other hand, ranks higher on these dimensions, which may lead individuals to perceive AI systems as posing a greater threat to womanhood. While largely mirroring Study 1, Study 2 examined not only attitudes toward feminists (to replicate findings of Study 1) but also the more general group of women. The adjustment was made because the lack of attitude changes toward feminists in Study 1 might not have quite represented the targeted discriminated group (women); indeed, feminists have been previously associated with negative perceptions representing a highly marginalized social group (e.g., Jenen et al., 2009).

### 3.1 Materials and Method

#### *3.1.1 Participants and Procedure*

A total of 350 participants from the US were recruited via Prolific. As in Study 1, the data of participants who did not identify as men or women were removed ($N$ = 8). The final sample included 342 participants ($M_{age}$ = 39.71, $SD$ = 12.61, 206 women). Similarly to Study 1, participants reported an average economic status ($M$ = 4.30, $SD$ = 1.35), adequate information about AI systems ($M$ = 4.84, $SD$ = 1.51), and average experience using AI systems in their daily lives ($M$ = 4.34, $SD$ = 1.96). The procedure of Study 2 was identical to that of Study 1.

#### *3.1.2 Materials*

Fake news reports. Three minor changes were made in the Study 1 news reports: an additional example of discrimination was included in the AI discriminator and human discriminator conditions, the human discriminators were changed from industry experts to "specialized officials," and the topic of the news report in the control condition was changed to the service industry. A team of researchers evaluated the texts to ensure clarity and consistency with the material from Study 1 (see Supplementary Materials).

Perceived threat to women. Due to the change in the topic of the news report for the control condition, the wording 'agriculture sector' was changed to 'service industry' in Study 2. The scales for AI systems (α = .92), specialized officials (α = .93), and the agriculture sector (α = .95) all demonstrated good reliability.

Collective action intentions. We added an item addressing collective action intentions for women's equal treatment through social media, which was adapted from previous research (e.g., Uluğ et al., 2023). Participants indicated their intentions to share and like social media content related to women's equal treatment (i.e., 'How likely are you to share and like related social media content for women's equal treatment?'), using the same 7-point scale (α = .90).

Attitudes toward women and feminists. In addition to measuring attitudes toward feminists, we used the same 6 bipolar statements to measure attitudes toward women, using a 7-point bipolar scale. Both attitudes toward women (α = .94) and feminists (α = .98) demonstrated good reliability.

### 3.2 Results

#### *3.2.1 Manipulation Check*

A Condition x Gender ANOVA revealed significant differences across the three conditions on the perceived discrimination against women ($F$(2, 336) = 72.79, $p$ < .001, $\eta_p^2$ = .302). Tukey's HSD tests showed that the perceived discrimination against women was stronger in the AI discriminator condition ($M$ = 5.24, $SD$ = 1.84) than in the control condition ($M$ = 3.11, $SD$ = 1.86, $p$ < .001), but not significantly different from the human discriminator condition ($M$ = 5.76, $SD$ = 1.44; $p$ = .060). Unlike Study 1, no main effect of Gender was observed ($p$ = .837). There were no significant interaction effects ($p$ = .994).

#### *3.2.2 Main Effects of Condition*

A MANOVA with all dependent variables found the multivariate effect of Condition ($F$(8, 674) = 9.10, $p$ < .001, $\eta^2$ = .097). As in Study 1, the univariate effect of Condition on perceived threat to women was significant ($F$(2, 339) = 30.16, $p$ < .001, $\eta_p^2$ = .151). Tukey's HSD tests showed that the perceived threat toward women was stronger in the human discriminator condition than in the AI discriminator ($p$ = .036) and control conditions ($p$ < .001). Similarly, the AI discriminator condition perceived higher threat to women than the control condition ($p$ < .001). We found no significant effect of Condition on collective action intentions for women's equal treatment ($p$ = .093), attitudes toward feminists ($p$ = .089), and attitudes toward women ($p$ = .168).

### *3.2.3 Exploratory Analyses*

To explore whether gender influenced perceived threat, collective action intentions, and attitudes toward women and feminists, another MANOVA was conducted with Gender and Condition as predictors. The multivariate effects of Condition ($F(8, 668) = 9.01$, $p < .001$, $\eta_p^2 = .097$), Gender ($F(4, 333) = 4.20$, $p = .002$, $\eta_p^2 = .048$), and the interaction was significant ($F(8, 668) = 2.02$, $p = .042$, $\eta_p^2 = .024$). Consistent with the initial MANOVA, the current analysis revealed a significant difference in perceived threat toward women across conditions ($F(2, 336) = 28.10$, $p < .001$, $\eta_p^2 = .143$). Similarly, the univariate effect of Condition on attitudes toward feminists was significant ($F(2, 336) = 3.38$, $p = .035$, $\eta_p^2 = .020$). However, Tukey's HSD tests did not reflect this univariate effect, as there were no significant differences across the three conditions in participants' attitudes toward feminists ($p$'s > .090). No significant univariate effect of Condition emerged for participants' collective action intentions for women's equal treatment ($p$ = .055), and their attitudes toward women ($p$ = .058).

Gender significantly affected participants' collective action intentions for women's equal treatment ($F(1, 336) = 13.27$, $p < .001$, $\eta_p^2 = .038$), and attitudes toward feminists ($F(1, 336) = 10.11$, $p = .002$, $\eta_p^2 = .029$). The means indicate that collective action intentions for women's equal treatment were higher for female participants ($M$ = 5.26, $SD$ = 1.62) than male participants ($M$ = 4.55, $SD$ = 1.87), and attitudes toward feminists showed the same pattern ($M_F = 5.29$, $SD_F = 1.64$; $M$ = 4.72, $SD$ = 1.73). Female and male participants did not differ in their perceived threat toward women ($p$ = .088), and attitudes toward women ($p$ = .053).

Moreover, an interaction between condition and gender predicted attitudes toward women ($F(2, 336) = 5.59$, $p = .004$, $\eta_p^2 = .032$). To investigate this interaction further, we conducted a follow-up univariate ANOVA on attitudes toward women. Tukey's HSD tests revealed that male participants in the human discriminator condition reported less favorable attitudes toward women ($M$ = 5.67, $SD$ = 1.25) than male participants in the control condition ($M$ = 6.37, $SD$ = .81, $p$ = .006), female participants in the human discriminator condition ($M$ = 6.39, $SD$ = .94, $p$ = .002), female participants in the AI discriminator condition ($M$ = 6.32, $SD$ = .85, $p$ = .005), and female participants in the control condition ($M$ = 6.25, $SD$ = 1.19, $p$ = .030). However, the difference was not significant between male participants in the human discriminator condition ($M$ = 5.67, $SD$ = 1.25) and male participants in the AI discriminator condition ($M$ = 6.27, $SD$ = .78, $p$ = .061). We found no significant interaction effects for perceived threat to women ($p$ = .429), collective action intentions for women's equal treatment ($p$ = .209), and attitudes toward feminists ($p$ = .056).

Lastly, we investigated whether perceived threat mediated the effect of condition (IV) on collective action intentions for women's equal treatment, attitudes toward feminists, and attitudes toward women (DV). Unlike Study 1, perceived threat did not mediate the effect of condition on collective action intentions for women's equal treatment ($B$ = .10, $SE$ = .06, 95% CI [-0.02, 0.23]), In contrast, perceived threat mediated the effect of condition on attitudes toward feminists ($B$ = .19, $SE$ = .06, 95% CI [0.07, 0.31]). The relative indirect effect of the human discriminator condition ($B$ = .29, $SE$ = .10, 95% CI [0.10, 0.49]) was stronger than that of the AI discriminator condition ($B$ = .19, $SE$ = .08, 95% CI [0.06, 0.37]), as indicated by the coefficients. Similarly, perceived threat mediated the effect of condition on attitudes toward women ($B$ = .09, $SE$ = .36, 95% CI [0.01, 0.16]). Again, the relative indirect effect of the human discriminator condition ($B$ = .13, $SE$ = .06, 95% CI [0.03, 0.25]) was stronger than that of the AI discriminator condition ($B$ = .09, $SE$ = .04, 95% CI [0.02, 0.18]).

## 3.3 Discussion

We replicated some findings from Study 1, but also uncovered some unique effects. First, Study 2 partially supported H1 and H2, demonstrating that both forms of discrimination trigger threats to women but AI discrimination was perceived as less threatening than human discrimination. These results align with Study 1 and the prior literature where AI systems were perceived as less threatening, presumably due to their lack of capacity for voluntary action (Bigman et al., 2023; Gray et al., 2007; Shank & DeSanti, 2018). Also replicating Study 1, no significant differences were found in collective action intentions for women's equal treatment and attitudes toward feminists and women (H3) in the hypothesized direction. Moreover, an exploratory MANOVA revealed a significant interaction between gender and condition, demonstrating that male participants who read about humans' discriminatory behaviors reported less favorable attitudes toward women, compared to male participants in the control condition and female participants. This pattern was not observed in Study 1 with Turkish participants and thus may be culturally specific, though further investigation is needed. This interaction can be explained by prior research demonstrating the advantaged group's willingness to oppose equality movements to 'protect' their privileges (e.g., Guvensoy et al., 2025; Van Laar et al., 2024). Lastly, in contrast to Study 1, the mediating effect of perceived threat emerged for attitudes but not for collective action intentions.

## 4. Discussion and Conclusions

The current studies examined whether and how discrimination by AI affects perceived threat toward the targeted group (women), attitudes toward the targeted and advocate groups (women and feminists), and collective action intentions for the targeted group. In two studies conducted in Türkiye and the US, we compared participants' responses to AI discrimination, human discrimination, and no discrimination. We hypothesized that, irrespective of the source, being exposed to gender discrimination would lead to increases in perceived threat to women, attitudes toward women and feminists, and motivation to engage in collective action for gender equality. We also expected that those reactions would be dampened when AI is the discriminator, given that AI is generally attributed less intention and agency, compared to humans (e.g., Bigman et al., 2023; Gray et al., 2007). Although attitudes and collective action intention remained unaffected, both discrimination conditions resulted in greater perceived threat to women, and discrimination was seen less threatening and harming when sourced by AI than humans.

These findings offer important implications regarding the potential risks of integrating AI into human life. While many recognize the benefits of AI for societal progress, its role in challenging – or reinforcing – existing hierarchical social structures has not received sufficient attention. Our studies demonstrated that discrimination was not seen as harmful when done by AI. Since discrimination by AI is not perceived as threatening as discrimination by humans, people may overlook the effects of its biased behaviors. In turn, AI systems may amplify existing unequal treatment of disadvantaged groups, and perhaps more critically, with such harmful outcomes insufficiently addressed, AI can normalize and further perpetuate social inequalities with limited accountability. Thus, it is critical for AI developers and policy makers to carefully consider the potential long-term risks of AI-driven discrimination.

On the other hand, collective action intentions for women's equal treatment and attitudes toward feminists and women did not differ across the three conditions. The findings may appear surprising because previous studies found reading about discrimination to result in positive intergroup outcomes (e.g., Bagci et al., 2017; Uluğ et al., 2023). Importantly, however, research also offers a more nuanced picture and shows that merely recognizing discrimination does not necessarily entail collective action intentions (Reimer et al., 2017), with recognition of one's privileges posing larger effects on willingness to act (e.g., Uluğ & Tropp, 2021). Research also highlights moral convictions and emotional reactions, like anger and sympathy, as key motivators of collective action in both advantaged (Iyer & Ryan, 2009; Van Zomeren et al., 2011) and disadvantaged groups (Iyer & Ryan, 2009; Van Zomeren et al., 2008, 2012) Likewise, perceived efficacy (i.e., the extent to which people believe in the impact of their actions) is another key motivator of collective actions (e.g., Van Zomeren et al., 2008). Taken together, we speculate that the specific scenario employed in the current work might not have been sufficient in eliciting emotional arousal or a sense of efficacy necessary to inspire action.

Notably, American male participants in the human discriminator condition reported less favorable attitudes toward women than male participants in the control condition and female participants. Research reports that advantaged groups show disinterest in collective action and support for the disadvantaged due to their zero-sum beliefs about equality (e.g., Van Laar et al. 2024). Closely imitating real-life circumstances, the human discriminator condition might have highlighted the differential treatment of men and women, leading enough male participants to show their willingness to maintain the favorable status quo. This pattern, however, was not observed in Türkiye. One speculative yet plausible account is that, compared to American men, Turkish men are less concerned about losing their privileges because the patriarchal system is perceived as rather stable (e.g., Bacchini et al., 2024).

Lastly, in both studies, male participants were in general less willing to engage in collective action for women's equal treatment and reported less favorable attitudes toward feminists. This gender difference aligns with existing findings in the literature that show women to be more motivated to defend their equality and feel more positively about feminists (e.g., Iyer & Ryan, 2009; Lemaster et al., 2015). While there were further differences across the two contexts in terms of how men and women responded to discrimination in general, these findings did not specifically relate to AI discrimination, therefore revealing that responses to AI discrimination were not substantially different across group status and country.

### 4.1 Limitations and Future Directions

Although the current studies offer unique contributions to literature by showing the potential effects of novel technologies on people's intergroup dynamics, its limitations must be acknowledged. First, although commonly used in research, self-report measures are susceptible to social desirability biases and thus might not have fully captured participants' genuine responses. For example, some participants may have felt uncomfortable reporting no intention to support women, especially during a time when femicides were being heavily discussed (Study 1). Participants may have reported stronger intentions to engage in collective action and positive attitudes than they actually intend or feel. Further research might be needed to use more implicit, as well as more behavioral measures to assess the outcome variables.

Another limitation concerns our limited understanding of why discrimination by AI versus humans resulted in different levels of threats toward women. AI is typically attributed less intention and agency than humans (e.g., Gray et al., 2007), but the present work did not directly test the mechanisms through which evaluations of human and AI perpetrators diverge. An alternative explanation, for example, is that AI discrimination was seen less consequential because it is assumed to be less prevalent than human discrimination. Further research is needed to scrutinize why AI discrimination was seen differently than human discrimination.

Finally, the political environment in Türkiye and the United States during data collection may have increased participants' baseline levels of collective action intentions. A month before the launch of Study 1, recent femicides were being addressed in Turkish media (Poyrazlar, 2024; We Will Stop Femicides Platform, 2024). Similarly, before the launch of Study 2, issues related to women's reproductive rights, including bans on women's emergency abortion access (Kruesi & Ungar, 2025; Lee, 2025), withholding payment for women's travel for abortion in the military (Kube, 2025) and defunding planned parenthood programs in multiple states (Chettiar, 2025; Ollstein, 2025) gathered attention. Some of this news, in addition to the unrest regarding the current government, increased public backlash, and led people to collective action for similar purposes (Deliso, 2025; King et al., 2025; Taylor, 2025). If participants were more prone to support collective action for gender equality at the time of the studies, the direct effects of discrimination on collective action intentions might have been less pronounced. Future research could replicate these studies in a less politically active climate or use different methods to manipulate discrimination, such as a simulated interaction between participants, an AI system, and a disadvantaged group member (i.e., target).

Overall, AI systems can reproduce biased and discriminatory outcomes that reinforce existing inequalities affecting disadvantaged groups. In this study, AI discriminators were perceived as less threatening than their human counterparts, yet neither increased collective action intentions nor improved attitudes toward the targeted group. These findings suggest that perceptions of discrimination may be obscured when harmful decisions are attributed to AI. Moreover, even when discrimination is recognized, awareness of discrimination alone is insufficient to mobilize support for social change. Further research is needed to examine how AI systems and their diverse forms of unfair treatment shape individual attitudes, behaviors, and social inequalities over the long term, and how society as a whole can promote awareness and effectively address these issues. Such work is crucial for understanding the cumulative societal consequences of discrimination in the age of AI and for informing ethical, regulatory, and policy responses.

## References

Allen, R., & Masters, D. (2020). Artificial intelligence: The right to protection from discrimination caused by algorithms, machine learning and automated decision-making. *ERA Forum*, *20*(4), 585–598. https://doi.org/10.1007/s12027-019-00582-w

Allport, G. W. (1954). *The nature of prejudice*. Addison-Wesley.

Allpress, J. A., Barlow, F. K., Brown, R., & Louis, W. R. (2010). Atoning for colonial injustices: Group-based shame and guilt motivate support for reparation. *International Journal of Conflict and Violence*, *4*(1), 75–88. https://doi.org/10.4119/IJCV-2816

Allpress, J. A., Brown, R., Giner-Sorolla, R., Deonna, J. A., & Teroni, F. (2014). Two faces of group-based shame: Moral shame and image shame differentially predict positive and negative orientations to ingroup wrongdoing. *Personality and Social Psychology Bulletin*, *40*(10), 1270–1284. https://doi.org/10.1177/0146167214540724

Bacchini, E. A. M., Scheepers, D., Ellemers, N., & Naber, M. (2024). Traditionally advantaged group members' affective and physiological responses to social change. *Social Psychology*, *55*(6), 295–305. https://doi.org/10.1027/1864-9335/a000567

Bagci, S. C., Turnuklu, A., & Tercan, M. (2020). Positive intergroup contact decreases the likelihood that prejudicial attitudes become avoidant behavioral tendencies. *European Journal of Social Psychology*, *50*(3), 597–613. https://doi.org/10.1002/ejsp.2646

Bagci, S. C., Verkuyten, M., Koc, Y., Turnuklu, A., Piyale, Z. E., & Bekmezci, E. (2020). Being tolerated and being discriminated against: Links to psychological well-being through threatened social identity needs. *European Journal of Social Psychology*, *50*(7), 1463–1477. https://doi.org/10.1002/ejsp.2699

Bagci, S. C., Çelebi, E., & Karaköse, S. (2017). Discrimination towards ethnic minorities: How does it relate to majority group members' outgroup attitudes and support for multiculturalism. *Social Justice Research*, *30*(1), 1-22.

Batson, C. D., Chang, J., Orr, R., & Rowland, J. (2002). Empathy, attitudes, and Action: Can feeling for a member of a stigmatized group motivate one to help the group? *Personality and Social Psychology Bulletin*, *28*(12), 1656–1666. https://doi.org/10.1177/014616702237647

Batson, C. D., Polycarpou, M. P., Harmon-Jones, E., Imhoff, H. J., Mitchener, E. C., Bednar, L. L., Klein, T. R., & Highberger, L. (1997). Empathy and attitudes: Can feeling for a member of a stigmatized group improve feelings toward the group? *Journal of Personality and Social Psychology*, *72*(1), 105–118. https://doi.org/10.1037/0022-3514.72.1.105

Becker, J. C., & Barreto, M. (2014). Ways to go: Men's and women's support for aggressive and nonaggressive confrontation of sexism as a function of gender identification. *Journal of Social Issues*, *70*(4), 668–686. https://doi.org/10.1111/josi.12085

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big?. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–623. https://doi.org/10.1145/3442188.3445922

Bigman, Y. E., Waytz, A., Alterovitz, R., & Gray, K. (2019). Holding robots responsible: The elements of machine morality. *Trends in Cognitive Sciences*, *23*(5), 365–368. https://doi.org/10.1016/j.tics.2019.02.008

Bigman, Y. E., Wilson, D., Arnestad, M. N., Waytz, A., & Gray, K. (2023). Algorithmic discrimination causes less moral outrage than human discrimination. *Journal of Experimental Psychology: General*, *152*(1), 4–27. https://doi.org/10.1037/xge0001250

Borrell, L. N., Jacobs, D. R., Williams, D. R., Pletcher, M. J., Houston, T. K., & Kiefe, C. I. (2007). Self-reported racial discrimination and substance use in the coronary artery risk development in adults study. *American Journal of Epidemiology*, *166*(9), 1068–1079. https://doi.org/10.1093/aje/kwm180

Branscombe, N. R., Schmitt, M. T., & Harvey, R. D. (1999). Perceiving pervasive discrimination among African Americans: Implications for group identification and well-being. *Journal of Personality and Social Psychology*, *77*(1), 135–149. https://doi.org/10.1037/0022-3514.77.1.135

Brown, N. D., Jacoby-Senghor, D. S., & Raymundo, I. (2022). If you rise, I fall: Equality is prevented by the misperception that it harms advantaged groups. *Science Advances*, *8*(18), eabm2385. https://doi.org/10.1126/sciadv.abm2385

Brown, R., & Cehajic, S. (2008). Dealing with the past and facing the future: Mediators of the effects of collective guilt and shame in Bosnia and Herzegovina. *European Journal of Social Psychology*, *38*(4), 669–684. https://doi.org/10.1002/ejsp.466

Caricati, L., & Owuamalam, C. K. (2020). System justification among the disadvantaged: A triadic social stratification perspective. *Frontiers in Psychology*, *11*, 40. https://doi.org/10.3389/fpsyg.2020.00040

Chettiar, I. (2025, February 5). Evidence that the right's assault on women is here. *Newsweek*. https://www.newsweek.com/evidence-that-rights-assault-women-here-opinion-2026061

Crawford, K. (2016, June 25). Artificial intelligence's white guy problem. *The New York Times*. https://www.nytimes.com/2016/06/26/opinion/sunday/artificial-intelligences-white-guy-problem.html

Danks, D., & London, A. J. (2017). Algorithmic bias in autonomous systems. *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence*, 4691–4697. https://doi.org/10.24963/ijcai.2017/654

Darling, K. (2012). Extending legal rights to social robots. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2044797

Darling, K. (2015). "Who's Johnny?" Anthropomorphic framing in human-robot interaction, Integration, and Policy. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2588669

Dastin, J. (2018, October 11). Insight – Amazon scraps secret AI recruiting tool that showed bias against women. *Reuters*. https://www.reuters.com/article/world/insight-amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK0AG/

Deliso, M. (2025, February 6). Protesters across the US rally against Trump administration policies. *ABC News*. https://abcnews.go.com/US/anti-trump-protests-nationwide/story?id=118501194

Dion, K. L., & Earn, B. M. (1975). The phenomenology of being a target of prejudice. *Journal of Personality and Social Psychology*, *32*(5), 944–950. https://doi.org/10.1037/0022-3514.32.5.944

Dovidio, J. F., Hewstone, M., Glick, P., & Esses, V. M. (2010). Prejudice, stereotyping and discrimination: Theoretical and empirical overview. In J. F. Dovidio, M. Hewstone, P. Glick, & V. M. Esses (Eds.), *The SAGE Handbook of Prejudice, Stereotyping and Discrimination* (pp. 3–28). SAGE Publications Ltd. https://doi.org/10.4135/9781446200919.n1

Ellemers, N., & Barreto, M. (2009). Collective action in modern times: How modern expressions of prejudice prevent collective action. *Journal of Social Issues*, *65*(4), 749–768. https://doi.org/10.1111/j.1540-4560.2009.01621.x

Ellemers, N., Van Knippenberg, A., & Wilke, H. (1990). The influence of permeability of group boundaries and stability of group status on strategies of individual mobility and social change. *British Journal of Social Psychology*, *29*(3), 233–

246. https://doi.org/10.1111/j.2044-8309.1990.tb00902.x

Eurostat. (2025). *Usage of AI technologies increasing in EU enterprises*. https://ec.europa.eu/eurostat/fr/web/products-eurostat-news/-/ddn-20250123-3#:~:text=In%202024%2C%2013.5%25%20of%20enterprises,growth%20from%208.0%25%20in%202023.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Ferrara, E. (2023). Fairness and bias in artificial intelligence: A brief survey of sources, impacts, and mitigation strategies. *Sci*, *6*(1), 3. https://doi.org/10.3390/sci6010003

Flood, M., Dragiewicz, M., & Pease, B. (2021). Resistance and backlash to gender equality. *Australian Journal of Social Issues*, *56*(3), 393–408. https://doi.org/10.1002/ajs4.137

Floridi, L., & Sanders, J. W. (2004). On the morality of artificial agents. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3848388

Geiselmann, R., Tsourgianni, A., Deroy, O., & Harris, L. T. (2023). Interacting with agents without a mind: The case for artificial agents. *Current Opinion in Behavioral Sciences*, *51*, 101282. https://doi.org/10.1016/j.cobeha.2023.101282

Giuliani, M., Mirnig, N., Stollnberger, G., Stadler, S., Buchner, R., & Tscheligi, M. (2015). Systematic analysis of video data from different human–robot interaction studies: A categorization of social signals during error situations. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00931

Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. *Science*, *315*(5812), 619–619. https://doi.org/10.1126/science.1134475

Gray, K., Yam, K. C., Eng, A., Dillion, D., & Waytz, A. (2025). The psychology of robots and artificial intelligence. In D. Gilbert, S. Fiske, E. Finkel, & W. Mendes (Eds.), *Handbook of social psychology 6th Edition* (6th ed.). Situational Press. https://doi.org/10.70400/IKVM1257

Hall, P., & Ellis, D. (2023). A systematic review of socio-technical gender bias in AI algorithms. *Online Information Review*, *47*(7), 1264–1279. https://doi.org/10.1108/OIR-08-2021-0452

Hayes, A. F. (2022). Introduction to mediation, moderation, and conditional process analysis: A regression-based approach (3rd ed.). The Guilford Press.

Hoskin, R. E., Thomas, E. F., & McGarty, C. (2019). Transnational contact and challenging global poverty: Intergroup contact intensifies (the right kind of) social identities to promote solidarity-based collective action for those low in social dominance. *Journal of Theoretical Social Psychology*, *3*(1), 23–34. https://doi.org/10.1002/jts5.24

Intahchomphoo, C., & Gundersen, O. E. (2020). Artificial intelligence and race: A systematic review. *Legal Information Management*, *20*(2), 74–84. https://doi.org/10.1017/S1472669620000183

Iyer, A., & Ryan, M. K. (2009). Why do men and women challenge gender discrimination in the workplace? The role of group status and in-group identification in predicting pathways to collective action. *Journal of Social Issues*, *65*(4), 791–814. https://doi.org/10.1111/j.1540-4560.2009.01625.x

Jenen, J., Winquist, J., Arkkelin, D., & Schuster, K. (2009). Implicit attitudes towards feminism. *Sex Roles*, *60*(1), 14-20. https://doi.org/10.1007/s11199-008-9514-3

Jonker, A., & Rogers, J. (2024, September 20). What is algorithmic bias? *IBM*. https://www.ibm.com/think/topics/algorithmic-bias

Kahn, P. H., & Shen, S. (2017). NOC NOC, who's there? A new ontological category (NOC) for social robots. In N. Budwig, E. Turiel, & P. D. Zelazo (Eds.), *New perspectives on human development* (1st ed., pp. 106–122). Cambridge University Press. https://doi.org/10.1017/CBO9781316282755.008

Kandiyoti, D. (2016). Locating the politics of gender: Patriarchy, neo-liberal governance and violence in Turkey. *Research and Policy on Turkey*, *1*(2), 103–118. https://doi.org/10.1080/23760818.2016.1201242

King, R. L., Moon, J., & Debushmann Jr, B. (2025, April 6). Anti-Trump protests held in cities across the US. *BBC*. https://www.bbc.com/news/articles/cz79ewg193ro

Knowles, E. D., & Lowery, B. S. (2012). Meritocracy, self-concerns, and whites' denial of racial inequity. *Self and Identity*, *11*(2), 202–222. https://doi.org/10.1080/15298868.2010.542015

Kruesi, K., & Ungar, L. (2025, March 6). What to know about the Idaho emergency abortion case dropped by the Trump administration. *PBS News*. https://www.pbs.org/newshour/health/what-to-know-about-the-idaho-emergency-abortion-case-dropped-by-the-trump-administration

Kube, C. (2025, January 31). Pentagon ends paying for travel for abortion and other reproductive care. *NBC News*. https://www.nbcnews.com/politics/national-security/pentagon-ends-policy-funding-travel-abortions-care-rcna190198

Kutlaca, M., Radke, H. R. M., & Becker, J. C. (2022). The impact of including advantaged groups in collective action against social inequality on politicized identification of observers from disadvantaged and advantaged groups. *Political Psychology*, *43*(2), 297–315. https://doi.org/10.1111/pops.12755

Landrine, H., & Klonoff, E. A. (1996). The schedule of racist events: A measure of racial discrimination and a study of its negative physical and mental health consequences. *Journal of Black Psychology*, *22*(2), 144–168. https://doi.org/10.1177/00957984960222002

Lee, C. (2025, March 7). Trump administration drops high-profile emergency abortion case, Leaving Advocates 'Devastated.' *TIME*. https://time.com/7265757/trump-lawsuit-abortion-idaho-ban-emergencies-activists-providers-respond/

Lemaster, P., Strough, J., Stoiko, R., & DiDonato, L. (2015). To have and to do: Masculine facets of gender predict men's and women's attitudes about gender equality among college students. *Psychology of Men & Masculinity*, *16*(2), 195–205. https://doi.org/10.1037/a0036429

Link, B. G., & Phelan, J. C. (2001). Conceptualizing stigma. *Annual Review of Sociology*, *27*(1), 363–385. https://doi.org/10.1146/annurev.soc.27.1.363

Lowery, B. S., Knowles, E. D., & Unzueta, M. M. (2007). Framing inequity safely: Whites' motivated perceptions of racial privilege. *Personality and Social Psychology Bulletin*, *33*(9), 1237–1250. https://doi.org/10.1177/0146167207303016

Maslej, N., Fattorini, L., Perrault, R., Gil, Y., Parli, V., Kariuki, N., Capstick, E., Reuel, A., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Niebles, J. C., Shoham, Y., Wald, R., Walsh, T., Hamrah, A., Santarlasci, L., … Oak, S. (2025). *Artificial intelligence index report 2025*. Stanford University Human Centered AI. https://hai.stanford.edu/ai-index/2025-ai-index-report

Mazzuca, S., Moscatelli, S., Menegatti, M., & Rubini, M. (2022). Men's reactions to gender inequality in the workplace: From relative deprivation on behalf of women to collective action. *Frontiers in Psychology*, *13*, 999750. https://doi.org/10.3389/fpsyg.2022.999750

McElheran, K., Li, J. F., Brynjolfsson, E., Kroff, Z., Dinlersoz, E., Foster, L., & Zolas, N. (2024). AI adoption in America: Who, what, and where. *Journal of Economics & Management Strategy*, *33*(2), 375–415. https://doi.org/10.1111/jems.12576

Miron, A. M., Warner, R. H., & Branscombe, N. R. (2011). Accounting for group differences in appraisals of social inequality: Differential injustice standards. *British Journal of Social Psychology*, *50*(2), 342–353. https://doi.org/10.1111/j.2044-8309.2010.02009.x

Moscatelli, S., Mazzuca, S., Menegatti, M., & Rubini, M. (2025). Women's participation in collective action for workplace gender equality: The role of perceived relative deprivation, resentment, and moral conviction. *Sex Roles*, *91*(4), 22. https://doi.org/10.1007/s11199-025-01573-7

Nadeem, A., Abedin, B., & Marjanovic, O. (2020). Gender bias in AI: A review of contributing factors and mitigating strategies. *Proceedings of the 31st Australasian Conference on Information Systems (ACIS 2020)*, *27*. https://aisel.aisnet.org/acis2020/27

Ollstein, A. M. (2025, March 31). Trump admin cuts tens of millions from Planned Parenthood. *Politico*. https://www.politico.com/news/2025/03/31/trump-admin-cuts-tens-of-millions-from-planned-parenthood-00261763

Osborne, D., & Sibley, C. G. (2015). Opposing paths to ideology: Group-based relative deprivation predicts conservatism through warmth toward ingroup and outgroup members. *Social Justice Research*, *28*(1), 27–51. https://doi.org/10.1007/s11211-014-0227-1

Osborne, D., Jost, J. T., Becker, J. C., Badaan, V., & Sibley, C. G. (2019). Protesting to challenge or defend the system? A system justification perspective on collective action. European Journal of Social Psychology, 49(2), 244-269.

Özdemir Sarıgil, B. (2022). Combating gender-based discrimination against women in Turkey: Achievements and the challenges ahead. In G. Yılmaz (Ed.), *Non-discrimination in Turkey* (pp. 59–77). Springer International Publishing. https://doi.org/10.1007/978-3-031-08399-0_4

Pascoe, E. A., & Smart Richman, L. (2009). Perceived discrimination and health: A meta-analytic review. *Psychological Bulletin*, *135*(4), 531–554. https://doi.org/10.1037/a0016059

Phillips, L. T., & Lowery, B. S. (2015). The hard-knock life? Whites claim hardships in response to racial inequity. *Journal of Experimental Social Psychology*, *61*, 12–18. https://doi.org/10.1016/j.jesp.2015.06.008

Poyrazlar, E. (2024, October 10). Outrage over murders of women forces Turkey's Erdoğan to harden laws. *Politico*. https://www.politico.eu/article/outrage-murder-women-turkey-erdogan-feminicide-laws-ak-party/

Reimer, N. K., Becker, J. C., Benz, A., Christ, O., Dhont, K., Klocke, U., Neji, S., Rychlowska, M., Schmid, K., & Hewstone, M. (2017). Intergroup contact and social change: Implications of negative and positive contact for collective action in advantaged and disadvantaged groups. *Personality and Social Psychology Bulletin*, *43*(1), 121–136. https://doi.org/10.1177/0146167216676478

Schmitt, M., & Maes, J. (2002). Stereotypic ingroup bias as self-defense against relative deprivation: Evidence from a longitudinal study of the German unification process. *European Journal of Social Psychology*, *32*(3), 309–326. https://doi.org/10.1002/ejsp.91

Schmitt, M. T., Branscombe, N. R., Postmes, T., & Garcia, A. (2014). The consequences of perceived discrimination for psychological well-being: A meta-analytic review. *Psychological Bulletin*, *140*(4), 921–948. https://doi.org/10.1037/a0035754

Schmitt, M. T., Spears, R., & Branscombe, N. R. (2003). Constructing a minority group identity out of shared rejection: The case of international students. *European Journal of Social Psychology*, *33*(1), 1–12. https://doi.org/10.1002/ejsp.131

Shank, D. B., & DeSanti, A. (2018). Attributions of morality and mind to artificial intelligence after real-world moral violations. *Computers in Human Behavior*, *86*, 401–411. https://doi.org/10.1016/j.chb.2018.05.014

Shank, D. B., North, M., Arnold, C., & Gamez, P. (2021). Can mind perception explain virtuous character judgments of artificial intelligence? *Technology, Mind, and Behavior*, *2*(3). https://doi.org/10.1037/tmb0000047

Sidanius, J., & Pratto, F. (1999). *Social dominance: An intergroup theory of social hierarchy and Oppression* (1st ed.). Cambridge University Press. https://doi.org/10.1017/CBO9781139175043

Swim, J. K., & Miller, D. L. (1999). White guilt: Its antecedents and consequences for attitudes toward affirmative action. *Personality and Social Psychology Bulletin*, *25*(4), 500–514. https://doi.org/10.1177/0146167299025004008

Taylor, A. (2025, February 18). 'Not My President's Day': Thousands gather in protest against Trump. *The Atlantic*. https://www.theatlantic.com/photo/2025/02/thousands-gather-protest-against-trump-musk/681721/

Tian, L., & Oviatt, S. (2021). A taxonomy of social errors in human-robot interaction. *ACM Transactions on Human-Robot Interaction*, *10*(2), 1–32. https://doi.org/10.1145/3439720

Tougas, F., & Veilleux, F. (1990). The response of men to affirmative action strategies for women: The study of a predictive model. *Canadian Journal of Behavioural Science / Revue Canadienne*

Uluğ, Ö. M., Chayinska, M., & Tropp, L. R. (2023). Does witnessing gender discrimination predict women's collective action intentions for gender justice? Examining the moderating role of perceived female support. *Journal of Community & Applied Social Psychology*, *33*(2), 501–518. https://doi.org/10.1002/casp.2642

Uluğ, Ö. M., & Tropp, L. R. (2021). Witnessing racial discrimination shapes collective action for racial justice: Enhancing awareness of privilege among advantaged groups. *Journal of Applied Social Psychology*, *51*(3), 248–261. https://doi.org/10.1111/jasp.12731

United Nations Development Programme. (2023). *2023 Gender social norms index (GSNI): Breaking down gender biases: Shifting social norms towards gender equality*. United Nations Development Programme. https://hdr.undp.org/content/2023-gender-social-norms-index-gsni#/indicies/GSNI

U.S. Equal Employment Opportunity Commission. (2022). *EEOC sues iTutorGroup for age discrimination*. U.S. Equal Employment Opportunity Commission. https://www.eeoc.gov/newsroom/eeoc-sues-itutorgroup-age-discrimination

Uysal, M. S., Uluğ, Ö. M., Kanık, B., & Aydemir, A. (2022). "The liberation of LGBTQ+ will also liberate heterosexuals": Heterosexual feminist women's participation in solidarity-based collective action for LGBTQ+ rights. *European Journal of Social Psychology*, *52*(2), 377–390. https://doi.org/10.1002/ejsp.2799

Van Laar, C., Van Rossum, A., Kosakowska-Berezecka, N., Bongiorno, R., & Block, K. (2024). MANdatory – Why men need (and are needed for) gender equality progress. *Frontiers in Psychology*, *15*, 1263313. https://doi.org/10.3389/fpsyg.2024.1263313

Van Zomeren, M., Postmes, T., & Spears, R. (2008). Toward an integrative social identity model of collective action: A quantitative research synthesis of three socio-psychological perspectives. *Psychological Bulletin*, *134*(4), 504–535. https://doi.org/10.1037/0033-2909.134.4.504

Van Zomeren, M., Postmes, T., & Spears, R. (2012). On conviction's collective consequences:

Integrating moral conviction with the social identity model of collective action. *British Journal of Social Psychology*, *51*(1), 52–71. https://doi.org/10.1111/j.2044-8309.2010.02000.x

Van Zomeren, M., Postmes, T., Spears, R., & Bettache, K. (2011). Can moral convictions motivate the advantaged to challenge social inequality?: Extending the social identity model of collective action. *Group Processes & Intergroup Relations*, *14*(5), 735–753. https://doi.org/10.1177/1368430210395637

Vargas-Solar, G. (2022). Intersectional study of the gender gap in STEM through the identification of missing datasets about women: A multisided problem. *Applied Sciences*, *12*(12), 5813. https://doi.org/10.3390/app12125813

Veilleux, F., & Tougas, F. (1989). Male acceptance of affirmative action for women: The results of alturistic or egoistical motives? *International Journal of Psychology*, *24*(1–5), 485–496. https://doi.org/10.1080/00207594.1989.10600060

Vescio, T. K., Sechrist, G. B., & Paolucci, M. P. (2003). Perspective taking and prejudice reduction: The mediational role of empathy arousal and situational attributions. *European Journal of Social Psychology*, *33*(4), 455–472. https://doi.org/10.1002/ejsp.163

Vlasceanu, M., & Amodio, D. M. (2022). Propagation of societal gender inequality by internet search algorithms. *Proceedings of the National Academy of Sciences*, *119*(29), e2204529119. https://doi.org/10.1073/pnas.2204529119

Wajcman, J. (2010). Feminist theories of technology. *Cambridge Journal of Economics*, *34*(1), 143–152. https://doi.org/10.1093/cje/ben057

We Will Stop Femicides Platform. (2024). *We Will Stop Femicides Platform October 2024 report*. https://kadincinayetlerinidurduracagiz.net/veriler/3126/we-will-stop-femicides-platform-october-2024-report

West, S. M., Whittaker, M., & Crawford, K. (2019). Discriminating systems: Gender, race, and power in AI. *AI Now Institute*. https://ainowinstitute.org/publications/discriminating-systems-gender-race-and-power-in-ai-2

Williams, D. R., Neighbors, H. W., & Jackson, J. S. (2003). Racial/Ethnic discrimination and health: Findings from community studies. *American Journal of Public Health*, *93*(2), 200–208. https://doi.org/10.2105/AJPH.93.2.200

World Economic Forum. (2024). *Global gender gap report 2024*. World Economic Forum. https://www.weforum.org/publications/global-gender-gap-report-2024/

World Economic Forum. (2025). *Global gender gap report 2025*. World Economic Forum. https://www.weforum.org/publications/global-gender-gap-report-2025/

Wright, S. C., Aron, A., McLaughlin-Volpe, T., & Ropp, S. A. (1997). The extended contact effect: Knowledge of cross-group friendships and prejudice. *Journal of Personality and Social Psychology*, *73*(1), 73–90. https://doi.org/10.1037/0022-3514.73.1.73

Wright, S. C., Taylor, D. M., & Moghaddam, F. M. (1990). Responding to membership in a disadvantaged group: From acceptance to collective protest. *Journal of Personality and Social Psychology*, *58*(6), 994–1003. https://doi.org/10.1037/0022-3514.58.6.994

Wright, S. C., & Tropp, L. R. (2002). Collective action in response to disadvantage: Intergroup perceptions, social identification, and social change. In I. Walker & H. J. Smith (Eds.), *Relative deprivation: Specification, development, and integration* (pp. 200–236). Cambridge University Press.