

ORIGINAL ARTICLE OPEN ACCESS

Feminist AI: Feminist Perspectives on Artificial Intelligence (AI) in the Workplace

Gendered Perceptions and the Use of Generative AI in Academia: An Experimental Vignette Study

Cinzia Calluso¹  | Valentina Iacopino²  | Anna Piazza³ ¹Department of Business and Management, LUISS University, Rome, Italy | ²Department of Economic Sciences and Business Management, Faculty of Economics, Università Cattolica del Sacro Cuore, Milan, Italy | ³School of Business and Operations of Strategy, University of Greenwich, London, UKCorrespondence: Cinzia Calluso (ccalluso@luiss.it)

Received: 31 October 2025 | Revised: 31 August 2026 | Accepted: 4 September 2026

Keywords: academia | experimental design | gender bias | generative AI | stereotypes

ABSTRACT

The adoption of Generative Artificial Intelligence (GenAI) is reshaping the very nature of work, yet its social meaning and evaluative implications—particularly for women—remain underexplored. Although existing research has mainly examined GenAI adoption, use, and algorithmic bias, less is known about how the use of GenAI shapes how individuals are evaluated by others. Drawing on the Stereotype Content Model, Role Congruity Theory and techno-feminism perspectives, this study examines how women's use of GenAI influences perceptions of competence and warmth in academia. Using a $2 \times 2 \times 2$ experimental vignette design [gender*GenAI adoption*task type (teaching vs. research)] with 651 participants, we examine how evaluations vary by academic gender, GenAI adoption, and task domain (teaching vs. research). We find that GenAI adoption generally enhances perceived competence. Women, however, gain the most in research tasks, where technical skills challenge gendered expectations. Conversely, GenAI use reduces perceived warmth, disproportionately penalizing women, especially in research. These findings reveal a “double bind,” highlighting how GenAI adoption may reproduce and intensify gendered evaluative tradeoffs within academic prestige hierarchies and the need for institutional strategies to mitigate relational penalties while recognizing competence.

1 | Introduction

The rapid diffusion of Generative Artificial Intelligence (GenAI) is transforming knowledge-intensive work (Storey et al. 2025), including academic work, as a paradigmatic context of knowledge production. Unlike previous waves of digitalization, GenAI directly affects the very core of many professions, blurring the boundaries between human expertise, knowledge production, and automated assistance. Yet this transformation cannot be understood only as a matter of productivity, work organization, or task redesign. GenAI is also imbued with social and organizational meanings. Its use may be interpreted as signaling changes in professional relations and in the distribution of technical expertise, power, politics, and status. These challenges are particularly salient in professional contexts, where professional

legitimacy also depends on the conscious and expert use of innovative technologies. In this sense, GenAI use becomes part of the evaluative repertoire through which social contexts assess individuals' work and conduct; its adoption may therefore shape how individuals' competence, credibility, and professional authority are evaluated (Orlikowski and Gash 1994). These implications are especially important because technologies can both challenge and reproduce existing inequalities in the workplace. Particular attention has already been paid to whether GenAI may help narrow gender gaps, especially given that women's use of technology has historically been associated with questions of legitimacy. Although GenAI is increasingly recognized as a new opportunity to reduce inequalities and, potentially, to lower the barriers that limit the full participation and the potential of

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2026 The Author(s). *Gender, Work & Organization* published by John Wiley & Sons Ltd.

underrepresented groups, it must also be acknowledged that women's use of innovative technologies—such as GenAI—continues to be associated with, and often reinforces, (gendered) judgments in professional contexts (Gross 2023). This risk is not simply due to fewer women using technology but rather stems from women's historically limited space and power to shape how technology is designed, positioned, valued, and evaluated. As a result, women's technological expertise is often subject to greater scrutiny, requiring them to repeatedly demonstrate competence and legitimacy, in line with the so-called “prove it again” bias (J. C. Williams et al. 2016, 2020). In investigating the potential impact of GenAI for women, it is crucial to acknowledge that technologies are not neutral, but rather socially determined artifacts (Compagni et al. 2015) designed and embedded within social constructs that inherently attribute gendered meanings to them. Feminist scholarship has largely contributed to understanding how this view of technology applies to digital tools, which are imbued with cultural codes and gendered meanings (Cockburn 1991, 1988, 1993, 1997, 2009, 2025; Faulkner 2001; Wajcman 2004). For instance, technologies designed for support and assistance are often described as feminized, whereas data coding, data analysis, and decision-making systems are frequently associated with masculine expertise or framed as technically neutral (Lohan and Wendy 2004; Wajcman 2004). This perceived neutrality is contested: while algorithmic systems may be designed to reduce some forms of human bias, their effects depend on how they are designed, trained, and embedded in organizational decision-making processes (Barocas and Selbst 2016; Fortin et al. 2021). However, GenAI use cannot be understood as a neutral individual choice. The decision to use it, the competence attributed to its users, and the value assigned to its adoption within organizations are interpreted within social contexts and may contribute to reproducing and legitimizing existing social orders. Despite these broader insights into the relationship between technology and social behavior, we still know remarkably little about how the use of GenAI influences perceptions of women in professional contexts. Existing evidence points to two related concerns: a persistent gender gap in the adoption and trust of digital tools (Tang et al. 2025), and the reproduction of stereotypes in their outcomes (Nivenitha and Nagarajan 2025). Hence, most discussions remain focused on technical aspects—such as how to train algorithms to minimize bias (Ali et al. 2026)—rather than on organizational or behavioral consequences of GenAI use. Yet, there remains a significant gap about how the use of GenAI shapes power dynamics, professional recognition, and legitimacy, particularly when adopted by women (Barry and Stephenson 2025; Nedungadi et al. 2024; Otis et al. 2024). This gap can be addressed by considering the fact that the use of GenAI may intersect with gender stereotypes through the dimensions of competence and warmth. On the one hand, GenAI adoption may be interpreted as a signal of innovation, technical ability, and enhanced competence. On the other hand, because technology use is often associated with automation, efficiency, and rationality, it may also raise concerns about authenticity, relational engagement, and warmth. This duality is particularly critical for women, as it generates tension between the competence signal associated with technology use and the stereotypically gendered expectations of warmth, care, and relational availability. We examine the existence and persistence of these dynamics in academia, a context characterized by continuous evaluations of competence and by a strong orientation

toward the attainment of status and professional legitimacy. In this sense, academia is not merely an empirical setting for our investigation but assumes specific theoretical relevance. Existing scholarship shows that academic evaluations are not neutral (Acker 1990, 1998; O'Connor 2020; van den Brink and Benschop 2012), and that women continue to be exposed to gendered and stereotyped judgments across different domains of academic work, including research, teaching, service, and grant acquisition (Guarino and Borden 2017; Herschberg et al. 2018; Järvinen and Mik-Meyer 2025a, 2025b). The distinctiveness of this context lies precisely in the socially shared meanings attached to these different task domains. In this study, we investigate the use of GenAI and its effect on perceptions of women academics. In this sense, the study opens the socio-evaluative black box of GenAI adoption and professional evaluation by examining the social mechanisms through which the same adoption cue is translated into judgments of competence and warmth. Specifically, we ask:

How does the use of Generative AI shape perceptions of competence and warmth, and how do these perceptions vary by gender and task domain?

We draw on the Stereotype Content Model (SCM; Fiske et al. 2002; Cuddy et al. 2008) and Role Congruity Theory (RCT; Eagly and Karau 2002) as our core theoretical frameworks, together with feminist technology studies (Faulkner 2001; Wajcman 2004), and research on gendered organizations in the higher education sector (Acker 1990; van den Brink and Benschop 2012; O'Connor 2020) to explore social meanings and evaluative implications of GenAI adoption. To address this question, we conduct a $2 \times 2 \times 2$ experimental vignette study, manipulating gender, GenAI adoption, and task domain, namely teaching versus research. Our findings show that GenAI use operates as a double-edged evaluative signal: it increases perceived competence, particularly for women in research-oriented tasks, while decreasing perceived warmth, with women facing distinctive penalties especially in research contexts. The study contributes to debates on how emerging technologies may reproduce, reshape, or intensify gendered evaluations in knowledge-intensive work.

2 | Theoretical Background and Hypotheses Development

2.1 | GenAI as a Socially Meaningful Cue of Competence and Warmth

In their modern meaning, technologies are often associated with efficiency, expertise, knowledge, and competence, and their use may therefore act as a signal of analytical and technical abilities. However, the way they are established and diffused within social contexts is neither neutral nor reducible to their functional use (Faulkner 2001; Wajcman 2004); rather, their implementation is imbued with socially shared meanings and associations. Technologies are, in fact, social artifacts that incorporate and elicit specific values and beliefs (Leonardi 2011; Orlikowski 1999). As Compagni et al. (2015) emphasize, technologies help define—and sometimes even modify—organizational and social resources, such as relationships and their characteristics. Their use therefore does not merely reflect a technical operation but also conveys social and value-based meanings that may influence how

adopters and nonadopters are perceived. This signaling function is particularly relevant in the case of GenAI, whose adoption may communicate not only technical proficiency, but also innovation and alignment with emerging professional norms and power relations (Kalluri 2020).

Moreover, the socially shared meanings associated with technologies are processed differently depending on individuals' cognitive schemas and evaluative frames. The SCM (Cuddy et al. 2008; Fiske et al. 2002; Connor and Fiske 2017; Fiske et al. 2018; Fiske et al. 2007; Fiske and Cuddy 2005) is particularly useful in this respect, as it describes how individuals and groups are socially evaluated. The model considers two core dimensions of social evaluation: (i) competence, which emphasizes ability, efficacy, and skill, and (ii) warmth, which captures relational traits, such as empathy, helpfulness, and cooperation. Although SCM has traditionally been applied to groups and individuals with distinctive characteristics (Kervyn et al. 2010), its application has more recently been extended to social artifacts and innovative technological entities such as AI, robots, and GenAI tools (McKee et al. 2023). Thus, the framework provides a useful lens for understanding how GenAI adoption is socially interpreted through the dimensions of competence and warmth.

When individuals use GenAI, others around them observe and evaluate that use. In doing so, observers do not simply judge the performance of the technology itself; they also evaluate the user through a technological lens. GenAI may therefore become a social cue through which others infer whether the user appears more competent, up to date, technically skilled or innovative, and whether this enhanced competence is associated with lower relational orientation or reduced authenticity. These judgments are highly consequential in professional contexts, precisely because professional legitimacy does not derive only from what individuals do, but also from how they are perceived by others (Compagni et al. 2015). This extension of evaluations of technology to evaluations of its users is critical, as attributes associated with technological artifacts may be transferred to those who adopt them through processes of social signaling (Berger et al. 1998; Podolny 2005). For example, technologies such as computers have historically been culturally associated with masculinity, rationality, and technical competence, and these associations may extend to their users.

GenAI users may therefore be evaluated through these socially shared associations. This corresponds to higher perceived competence and greater consistency with the agentic dimension of the SCM model. We therefore expect that the adoption of GenAI is associated with higher perceived competence, innovation and productivity. Simultaneously, increasing and continuous interaction with machines may produce perceptions of relational distance, reduced humanity, and—consistent with the SCM framework—lower warmth (Fiske et al. 2002). In organizations, warmth translates into trust, cooperation, and a community-oriented mindset (Kervyn et al. 2010). Because the use of advanced technologies can be associated with efficiency, distance, and reduced authenticity, the use of GenAI may be interpreted as implying lower relational engagement. If GenAI is interpreted as a social cue associated with technical expertise, innovation, and efficiency, its adoption may decrease perceived

warmth by making users appear less relationally engaged, empathic or socially connected. This may translate into weaker judgments of users' authenticity and emotional availability. More formally:

H1a. *The use of GenAI will be positively associated with perceived competence.*

H1b. *The use of GenAI will be negatively associated with perceived warmth.*

2.2 | Technology, Gender, and Stereotypes

Techno-feminist scholarship has long argued that technologies are not neutral artifacts within organizations. Rather, their meanings are produced and sustained through broader relations of power, legitimacy, and gender (Cockburn 1988, 1997, 2009, 2025; Haraway 1991; Wajcman 1991, 2004; Faulkner 2001). Technologies do not play a merely functional role, but operate as socially constructed, politically situated, and socially meaningful artifacts that contribute to defining, shaping, and legitimizing particular social orders. They may reproduce gendered assumptions embedded in existing social arrangements, but they may also become sites through which such assumptions are contested and renegotiated (Haraway 1991; Wajcman 2004). This broader literature explains how gendered meanings become culturally associated with technologies.

This argument has been extended to disruptive technologies, such as GenAI and data-driven systems (D'Ignazio and Klein 2020; Browne et al. 2023; Toupin 2024; Wajcman and Young 2023; West et al. 2019; Noble 2018). Although often framed through the language of technical neutrality, digital and AI-based technologies are culturally associated with traits of rationality, technical mastery, control, and expertise, all historically coded as masculine (Lohan and Wendy 2004; Wajcman 2004; Cheryan et al. 2015). GenAI may therefore be socially interpreted as a masculine-coded technology, insofar as its use is associated with domains such as computer science, data analysis, programming, and algorithmic decision-making. Even practices such as prompting may be read as evidence of analytical ability, technical control, and rational problem-solving. In this sense, GenAI does not simply enter organizations as a neutral tool, but as a socially meaningful evaluative cue that may be interpreted through existing gendered roles and hierarchies.

When women use these technologies, gender stereotypes and technological meanings may come into tension. On the one hand, as argued above, the use of GenAI may signal competence, technical skill, innovation, and alignment with agentic forms of professional evaluation. On the other hand, gender stereotypes traditionally associate women with communal traits such as warmth, empathy, care, collaboration, and relational engagement. This creates a potential double bind (Jamieson 1995): women who appear assertive and technologically skilled may be judged as cold or unfeminine, whereas those who conform to communal expectations may be perceived as less suitable for leadership, innovation, or technological authority (Rudman 1998; Rudman and Glick 2001).

Although the SCM allows us to identify competence and warmth as the main evaluative dimensions through which

GenAI users are judged, RCT (Eagly and Karau 2002) helps explain why these judgments may vary by gender. RCT suggests that social evaluations are shaped by the perceived congruity or incongruity between gendered role expectations and the specific practices, behaviors, or domains in which individuals are observed. In the case of GenAI, women's adoption of a technology associated with rationality, control, and technical expertise may reduce perceived role incongruity on the competence dimension, allowing them to be seen as more technically skilled and professionally capable. For men, by contrast, GenAI adoption is more consistent with agentic gender expectations and may therefore produce weaker competence gains. At the same time, women's use of GenAI may increase incongruity on the warmth dimension, because it may distance them from communal expectations of care, empathy, and relational engagement. In this sense, SCM identifies the evaluative dimensions through which GenAI users are judged, whereas RCT clarifies why the same technological signal may produce different evaluative consequences for women and men. We therefore expect that women will gain more than men in perceived competence from GenAI adoption but will also face a greater penalty in perceived warmth. More formally:

H2a. *Gender will moderate the relationship between GenAI use and perceived competence, such that the increase in competence associated with GenAI use will be stronger for women than for men.*

H2b. *Gender will moderate the relationship between GenAI use and perceived warmth, such that the reduction in warmth associated with GenAI use will be stronger for women than for men.*

2.3 | Gendered Evaluations of GenAI Use Across Academic Domains

Academia constitutes a theoretically relevant organizational field in which competence, legitimacy, authority, and relationality are continuously evaluated. Drawing on feminist organizational theory, universities can be understood not as neutral communities, but as organizational contexts where relational dynamics, hierarchies, and, more broadly, the entire social order may reproduce gendered assumptions and stereotypes (Benschop and Brouns 2003; Acker 1990, 1998). Although academia is often portrayed as a meritocratic field guided by values such as transparency, ethics, excellence, rigor, social relevance, and international standing, prior research has shown that these principles may coexist with the reproduction and perpetuation of gender inequalities (Araneda-Guirriman et al. 2023; van den Brink and Benschop 2012; O'Connor 2020). The very concept of excellence, for example, does not refer to a purely objective standard, but can operate as a gender-sensitive evaluative dimension through which individuals' professionalism can be recognized or contested (Herschberg et al. 2018; Lagesen and Suboticki 2024; van den Brink and Benschop 2012). As a result, socially shared meanings associated with academic work, such as innovation and technical competence, are attributed to men and women based on gendered evaluative schemas. Consequently, women may be less readily recognized as embodying technical competence and authority (Phillips et al. 2022) and more readily associated with support, service, and relational work.

In this context, competence and warmth—as well as the agentic and communal traits associated with different tasks—operate as evaluative lenses through which academic domains and scholars' professional legitimacy are judged (Fiske et al. 2002). For example, research is associated with typically agentic meanings such as excellence, rigor, originality, prestige, intellectual depth, funding acquisition, international recognition, and career advancement. Teaching and service activities, by contrast, are more often coded as communal and associated with availability, support, care, and orientation toward relationships and students. These communal expectations are not gender-neutral: women faculty are frequently cast in explicitly caring roles and expected to embody care and relational availability, particularly in teaching (Blalock and Stefanese-Yates 2024). These domains are therefore neither equivalent in terms of academic prestige nor neutral with respect to gender.

To fully understand the organizational dynamics these categories may generate, it is important to integrate them within RCT (Eagly and Karau 2002). Prior research has shown that women are often associated with communal domains such as teaching, mentoring, and service (Guarino and Borden 2017; Heijstra et al. 2017; Järvinen and Mik-Meyer 2025a, 2025b), whereas agentic tasks such as research and grant acquisition are more easily aligned with masculine behaviors and identities.

In the context of AI adoption, this suggests that the way GenAI use modifies individuals' perceived competence may vary depending on the domain in which it takes place. In the case of competence, we expect that, in agentic tasks such as research, the use of GenAI may reduce the perceived incongruity between feminine role expectations and the agentic demands associated with effectiveness, performance, and technical expertise. As a result, women who use GenAI in research-oriented activities may be perceived as more competent. In the case of warmth, by contrast, we expect women who use GenAI in agentic domains to be particularly penalized, in line with research on prescriptive gender stereotypes and backlash toward agentic women (Rudman and Glick 2001). For instance, in research-oriented tasks, the use of a masculine-coded technology may compound the perceived violation of communal expectations traditionally associated with women, as they are evaluated within a domain already coded as masculine and associated with authority, excellence, and competence. More formally:

H3a. *The joint effect of GenAI use and gender on perceived competence will vary as a function of task type, with women being perceived as more competent in agentic (i.e., research) than in communal tasks (i.e., teaching).*

H3b. *The joint effect of GenAI use and gender on perceived warmth will vary as a function of task type, with women being more penalized in agentic (i.e., research) than in communal (i.e., teaching) tasks.*

3 | Materials and Methods

3.1 | Setting

Our study is situated within academia, the organizational context in which academics' work is continuously assessed by peers, students, institutions, and external audiences. Academic

activity is also structured around institutionally distinct task domains, including research and teaching, which provide the empirical basis for our task manipulation. As theorized above, academia is particularly suitable for this study because evaluation is structurally embedded in everyday academic work. Hiring, promotion, peer review, grant allocation, and student feedback all produce recurrent judgments of competence, legitimacy, and professional worth. These judgments are not neutral: they unfold within gendered evaluative hierarchies. In this context, GenAI use may become an additional evaluative cue through which academics are assessed. Yet, existing studies on GenAI in academia have mainly examined adoption practices, perceptions, and barriers among researchers themselves (Andersen et al. 2025; Dorta-González et al. 2024; Hsiao and Tang 2025), leaving open the question of how academics who use these technologies are perceived by others, and whether such perceptions vary by gender and task domain. Academia therefore provides a theoretically meaningful and empirically tractable context in which to examine how GenAI use shapes perceptions of competence and warmth, and their interaction with gender and task domain.

3.2 | Design

The study employs a $2 \times 2 \times 2$ between-subjects factorial design, manipulating gender_{vignette} (female vs. male), AI adoption (adoption vs. nonadoption), and task (teaching vs. research), leading to eight possible experimental conditions. Because the study focuses on the social meaning of AI adoption, we use GenAI as an umbrella term encompassing widely available generative AI tools—that is, ChatGPT, Claude, Gemini, and similar systems—typically employed for activities such as brainstorming, text editing, or supporting information analysis. Our aim is not to compare the technical affordances or normative implications of specific GenAI applications. Rather, the experimental manipulation isolates the social meaning attached to GenAI adoption itself, allowing us to examine how this cue informs evaluations of the user across gender and task domains. To this aim, we employed a vignette-based experimental design, a methodological approach particularly suited to the study of evaluative perceptions, which are inherently subjective and context-dependent. By combining controlled manipulation with realistic scenarios, vignettes allow for causal inference while preserving ecological validity (Aguinis and Bradley 2014; Atzmüller and Steiner 2010; Highhouse 2009).

The study relies on primary data collected through an online, self-administered survey, implemented and distributed via Qualtrics. Online administration is consistent with established practice in experimental vignette research, as it allows access to geographically dispersed participants, facilitates random assignment to experimental conditions, and ensures consistency in stimulus presentation (Aguinis and Bradley 2014; Atzmüller and Steiner 2010). Moreover, self-administered formats are particularly appropriate for studying social evaluation, as they reduce interviewer effects and social desirability bias, enabling respondents to express their judgments more freely (Podsakoff et al. 2003). The survey was organized into three sections. The first section was devoted to collecting demographic and professional information, including age, gender, marital status, parental

status, migrant status, race, ethnicity, job position, and institution type (teaching/research and public/private). These variables were included as controls to account for potential sources of heterogeneity in evaluative perceptions. Prior research shows that judgments of competence and warmth are systematically shaped by social characteristics such as gender, ethnicity, and status-related attributes. For instance, gender has been shown to influence standards of competence evaluation (Biernat and Fügen 2001), whereas ethnicity and gender jointly affect perceptions of task competence (Foschi and Buchan 1990). In addition, status-related characteristics—such as those proxied by organizational role—have been found to shape the relationship between competence and warmth in social evaluation (Lin et al. 2011). In this vein, demographic and professional characteristics such as age, gender, organizational role, institutional context, and background variables (e.g., migration status, race, and ethnicity) may influence how individuals interpret and evaluate signals related to competence, relationality, and technology use. Including these variables therefore allows us to more precisely estimate the effects of the experimental manipulations. In the second section, participants were randomly assigned to one of the eight vignettes, ensuring experimental control through random allocation to conditions and systematic manipulation of independent variables. The vignettes employed in the study are available in Supporting Information S1: Appendix A. The third section assessed perceived warmth and competence. In line with the SCT literature (e.g., Cuddy et al. 2009; Fiske and Cuddy 2005; Krings et al. 2011), competence and warmth were operationalized through sets of trait descriptors rather than through a single validated psychometric scale. Participants were asked to evaluate the target described in the vignette on a series of attributes commonly used to capture perceived competence and warmth, using a 7-point Likert scale. Specifically, competence was measured using six items—competent, confident, capable, efficient, intelligent, skillful—and warmth was measured using five items—friendly, well-intentioned, trustworthy, good-natured, and sincere.

3.3 | Participants

The study targets individuals operating within academic contexts, as the scenarios presented in the vignettes refer to academic work and evaluation processes. 955 participants took part in the study; 304 participants were excluded due to revoked consent (i.e., participants did not complete the full experiment). The final sample includes 651 participants (341 male, 296 female, 2 nonbinary, 12 prefer not to say; age = 43.75 ± 12.69) with roughly 81 participants allocated to each experimental condition (Teaching-noAI-Male = 81; Teaching-noAI-Female = 84; Teaching-AI-Male = 79; Teaching-AI-Female = 84; Research-noAI-Male = 83; Research-noAI-Female = 77; Research-AI-Female = 83; Research-AI-Male = 80).

In line with the Declaration of Helsinki and APA ethical standards, participants were asked to provide their informed consent prior to participation. Further, all data was collected anonymously and analyzed in aggregate. Participants were fully informed about the study and their right to withdraw at any time.

Participants were recruited by means of various channels. Specifically, the sampling strategy combined purposeful and snowball approaches targeting academic participants. All the authors posted an invitation to participate via their personal social media (e.g., LinkedIn, academic groups on Facebook etc.), as well as through direct and indirect academic contacts (e.g., inviting colleagues to participate and encouraging them to share the invitation within their professional networks, consistent with a snowball sampling approach). Finally, to maximize reach and enhance the heterogeneity of the sample within academia, a series of invitation emails were sent to publicly available email addresses (obtained by universities and research institutions websites), for a total of approximately 17,000 emails distributed. This recruitment strategy aimed to capture a broad and heterogeneous sample of academics across institutions and contexts, rather than achieving statistical representativeness, consistent with experimental research designs that prioritize internal validity and causal inference through controlled manipulation over population generalizability (Shadish et al. 2002; Calder et al. 1981; Highhouse 2009). Table 1 provides an overview of the demographic and work-related characteristics of the sample.

3.4 | Statistical Analyses

3.4.1 | Hypotheses Testing

The first set of analysis was computed to evaluate the effect of GenAI on the perception of competence and warmth of men and women in academia across teaching and research duties. The statistical significance of such effects was tested using two separate analyses computed using General Linear Models (GLM).

The first model was run using competence ratings as the dependent variable, and with $gender_{vignette}$ (M, F), AI adoption (AI, noAI), task type (teaching, research), and all second- and third-level interactions, as fixed effects. The model also included a series of covariates, namely age, $gender_{participant}$, marital status, parental status, migrant status, race, ethnicity, job position, and institution type (teaching/research and public/private).

The second model was identical to the first but used warmth ratings as the dependent variable. All the post hoc inspections of the interaction effects were corrected for multiple testing using the Fisher's Least Significant Difference (LSD) test (L. J. Williams and Hervé 2010).

3.4.2 | Exploratory Analysis

An additional set of GLMs were performed to explore the gendered effect in perceived warmth and competence, as a function of participants' gender. We estimated two separate models—one for competence rating, and one for warmth ratings. Both models followed the same specification described in the previous section, with the addition of the main effect of the factor $gender_{participant}$ (male, female), along with all the second-, third-, and fourth-level interactions. Full details of the analysis are provided in the Supporting Information S1: Appendix C.

4 | Results

4.1 | Hypotheses Testing

The results indicated that the first model, with competence scores as the dependent variable, was significant overall ($F = 3.46$, $p < 0.001$). Specifically, the main effect of AI adoption was significant ($F = 22.52$, $p < 0.001$), indicating that AI adoption was associated with higher perceived competence compared to non-adoption, thus confirming hypothesis H1a. The main effects of task ($F = 1.06$, $p = 0.30$) and $gender_{vignette}$ ($F = 2.75$, $p = 0.10$) were nonsignificant as were the two-way interactions (task*AI adoption: $F = 0.49$, $p = 0.48$; task* $gender_{vignette}$: $F = 0.15$, $p = 0.70$; AI adoption* $gender_{vignette}$: $F = 0.51$, $p = 0.47$). The lack of statistical significance of the AI adoption by $gender_{vignette}$ interaction indicates that H2a was not supported at the level of the hypothesized interaction effect. Importantly, the three-way interaction effect task* $gender_{vignette}$ *AI adoption was significant ($F = 6.49$, $p < 0.01$). The post hoc inspection of the third-level interaction effect revealed that in teaching activities, AI use increases the perceived competence of academics both men and women, though the effect is stronger for males (no-AI-M < AI-M: $p < 0.001$) than females (no-AI-F < AI-F: $p = 0.12$). In the no-AI condition, the difference in perceived competence between male and female is significant (no-AI-M < no-AI-F: $p < 0.05$), penalizing males more heavily; with AI, the gap is not significant (AI-M = AI-F: $p = 0.43$). In research activities, AI increases the perceived competence of females (no-AI-F < AI-F: $p < 0.01$) but not males (no-AI-M = AI-M: $p = 0.17$). With AI, females are rated marginally more competent than males (AI-F > AI-M: $p = 0.09$), whereas without AI, the difference is not significant (no-AI-F = no-AI-M: $p = 0.23$). The comparisons across research and teaching activities, with or without AI, across males and females, revealed that most of the comparisons of interest were nonsignificant, with the only exception of the comparison between research and teaching without AI of male vignettes, which resulted in a marginally significant effect (T-no-AI-M < R-no-AI-M: $p = 0.08$), which suggests that men who do not use AI are more heavily penalized when they do so in teaching rather than in research tasks. No such effect was found for women (T-no-AI-F = R-no-AI-F: $p = 0.89$). Supporting Information S1: Table S1 (Appendix B) provides an overview of the post hoc comparisons. Overall, the significance of the third-level interaction, along with the results of the post hoc tests, supports hypothesis H3a, showing that women, indeed, gain more competence than men when using AI, specifically in agentic tasks such as research.

Among covariates, only age was significant ($p < 0.001$), showing that as participants' age increased, perceived competence tended to decrease, resulting in more severe evaluations, regardless of the experimental condition. Table 2 provides a comprehensive report of the results of the GLM computed on competence ratings, and Figure 1 displays the graphical representation of the effects.

In the second model, with warmth ratings as the dependent variable, the overall model was significant ($F = 2.39$, $p < 0.001$). In particular, the main effect of task was significant ($F = 14.11$, $p < 0.001$), showing that teaching activities are perceived as warmer than research activities, regardless of gender or AI

TABLE 1 | Demographic and work-related characteristics of the sample.

Demographics			
Marital status		Race	
Married	49.92%	White	78.34%
In a relationship	26.57%	Prefer not to say	7.22%
Single	20.12%	Asian or Pacific Islander	6.61%
Separated/divorced	3.38%	Hispanic or latino/a/x	4.45%
Parental status		Middle eastern or North African	2.30%
Nonparent	52.69%	Black	0.77%
Parent	47.31%	Indigenous	0.31%
Geographical area (nationality)		Ethnic origin	
Europe	83.41%	Western Europe	80.95%
North America	10.08%	Eastern Europe	4.92%
Asia	7.32%	Prefer not to say	4.76%
South America	1.63%	South and Southeast Asia	3.38%
Oceania	1.46%	East and Central Asia	2.15%
Prefer not to say	1.30%	West Asia/Middle East	1.54%
Africa	0.65%	South America	0.92%
Migrant status ^a		Sub-Saharan Africa	0.77%
Yes	80.34%	North Africa	0.31%
No	17.97%	North America	0.31%
Prefer not to say	1.69%		
Work related information			
Position		Public/private institution	
Associate professor	24.12%	Public	81.57%
Full professor	19.05%	Private	18.43%
Assistant professor	17.97%	Type of institution	
PhD student	16.59%	Research	61.75%
Postdoctoral fellow	11.83%	Teaching	38.25%
Lecturer	7.37%		
Research assistant	2.61%		
Teaching assistant	0.46%		

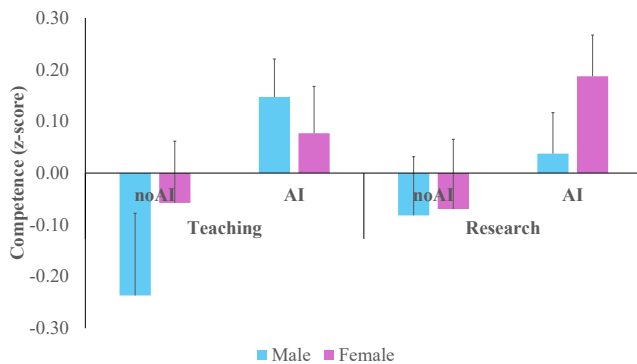
^aMigrant: refers to people who work in a country other than the one of their nationality.

adoption. AI adoption also had a significant main effect ($F = 8.62, p < 0.001$), being associated with lower perceived warmth across tasks and genders; this result confirms hypothesis H1b. The main effect of gender_{vignette} ($F = 0.68, p = 0.41$) and the two-way interactions (Task*AI adoption: $F = 0.16, p = 0.69$; Task*Gender_{vignette}: $F = 1.65, p = 0.20$; AI adoption*Gender_{vignette}: $F = 1.03, p = 0.31$) were nonsignificant. The nonsignificant AI adoption*Gender_{vignette} interaction indicates that H2b was not supported at the level of the hypothesized interaction effect. The three-way interaction ($F = 0.15, p = 0.70$) was also found to be nonsignificant; nonetheless, post hoc inspection revealed that although AI decreased perceived warmth overall, this penalty was significant only for women across both teaching (no-AI-F > AI-F: $p < 0.05$) and research (no-AI-F > AI-F: $p = 0.05$). Although the AI adoption*Gender_{vignette}

interaction did not support H2b, the post hoc comparisons revealed a pattern consistent with the underlying theoretical expectation, as the reduction in perceived warmth associated with AI use was significant for women in both research and teaching, but not for men. The comparisons between teaching and research task also revealed that women are penalized more in their perceived warmth for using AI when employed for research as compared to teaching tasks (R-AI-F < T-AI-F: $p < 0.05$). Further, the same comparison across the no-AI condition, revealed that when not using AI women are perceived as warmer in teaching as compared to when they do research (R-no-AI-F < T-no-AI-F: $p < 0.01$); in other words, even when not using AI, women are perceived as less warm when they do research as compared to teaching and they are penalized even more when they use AI. Although the

TABLE 2 | Results of the GLM performed using the competence ratings as a dependent variable.

	Type III sum of squares	Mean square	<i>F</i>	<i>p</i>
Corrected model	62.25	3.46	3.48***	0.00
Task	1.05	1.05	1.06	0.30
AI adoption	22.37	22.37	22.52***	0.00
Gender _{vignette}	2.73	2.73	2.75	0.10
Task * AI adoption	0.49	0.49	0.49	0.48
Task * gender _{vignette}	0.15	0.15	0.15	0.70
AI adoption * gender _{vignette}	0.51	0.51	0.51	0.47
Task * AI adoption * gender _{vignette}	6.45	6.45	6.49**	0.01
Gender _{participant}	1.16	1.16	1.17	0.28
Age	12.43	12.43	12.51***	0.00
Marital status	2.54	2.54	2.55	0.11
Parental status	1.22	1.22	1.23	0.27
Geographical area	0.00	0.00	0.00	0.95
Migrant status	0.22	0.22	0.22	0.64
Race	0.32	0.32	0.32	0.57
Ethnic origin	0.12	0.12	0.12	0.73
Job position	1.49	1.49	1.50	0.22
Public/private institute	2.63	2.63	2.64	0.10
Type of institute (teaching, research)	1.55	1.55	1.56	0.21

****p* < 0.001.***p* < 0.01.**FIGURE 1** | Bar-plots of the results of the GLM performed using the competence ratings as a dependent variable.

nonsignificant three-way interaction did not formally support H3b, the post hoc comparisons revealed a pattern consistent with the hypothesis, showing that women were more penalized in agentic than in communal tasks. This pattern suggests a double punishment: women are perceived as losing warmth both for engaging in research and for employing AI in their activities. Supporting Information S1: (Table S2, Appendix B) provides an overview of the post hoc comparisons.

Among the covariates, only age was significant ($p < 0.05$), showing that as participants' age increased, perceived warmth tended to decrease, resulting in more severe evaluations, regardless on the experimental condition. The effect of gender_{participant} was found marginally significant ($p = 0.09$), indicating that women tend to evaluate the academics described in the vignettes as less warm, compared to male respondents,

regardless the experimental condition. Table 3 provides a comprehensive report of the results of the GLM computed on competence ratings, and Figure 2 displays the graphical representation of the effects.

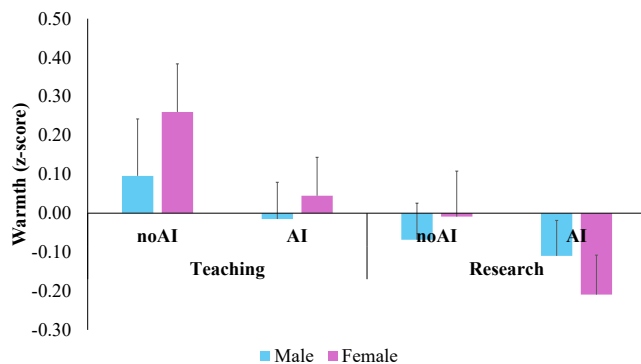
4.2 | Exploratory Analysis

In an exploratory follow-up analysis, we examined whether competence and warmth evaluations were influenced by participants' gender and its interaction with task type, AI adoption, and vignette gender. Full results are reported in the Supporting Information S1: (Appendix C).

The results for competence ratings replicated the main findings, showing overall higher competence ratings in the AI condition. Moreover, competence evaluations varied as a function of task type, vignette gender, and participant gender. Men tended to evaluate the male vignette in teaching tasks less positively than women evaluating female vignette involved in research tasks, reflecting a gendered evaluation pattern consistent with traditional gender roles, given that teaching is a communal task. Women also showed a marginal tendency to rate female academics in research more favorably than male ones, suggesting a potential women solidarity effect. A similar marginal pattern emerged for the interaction between AI adoption, vignette gender, and participant gender: in the no-AI condition, women tended to rate female academics more positively than male academics, whereas men showed no such difference. Hence, women seem to apply a gendered evaluation of competence, with male academics being penalized more for not adopting AI—as might be expected based on gender role norms—compared to female academics.

TABLE 3 | Results of the GLM performed using the warmth ratings as a dependent variable.

	Type III sum of squares	Mean square	<i>F</i>	<i>p</i>
Corrected model	43.00	2.39	2.40***	0.00
Task	14.06	14.06	14.11***	0.00
AI adoption	8.59	8.59	8.62***	0.00
Gender _{vignette}	0.68	0.68	0.68	0.41
Task * AI adoption	0.16	0.16	0.16	0.69
Task * gender _{vignette}	1.64	1.64	1.65	0.20
AI adoption * gender _{vignette}	1.03	1.03	1.03	0.31
Task * AI adoption * gender _{vignette}	0.15	0.15	0.15	0.70
Gender _{participant}	2.86	2.86	2.87†	0.09
Age	3.80	3.80	3.81†	0.05
Marital status	1.42	1.42	1.42	0.23
Parental status	0.11	0.11	0.11	0.74
Geographical area	1.61	1.61	1.62	0.20
Migrant status	0.07	0.07	0.07	0.79
Race	2.49	2.49	2.50	0.11
Ethnic origin	1.05	1.05	1.05	0.31
Job position	2.77	2.77	2.78	0.10
Public/private institute	1.64	1.64	1.65	0.20
Type of institute	0.21	0.21	0.21	0.65

****p* < 0.001.†*p* < 0.1.**FIGURE 2** | Bar-plots of the results of the GLM performed using the warmth ratings as a dependent variable.

In a second exploratory analysis, we examined whether warmth evaluations were influenced by participants' gender, task type, and AI adoption. The overall pattern replicated the main findings: warmth was rated higher for teaching tasks than for research tasks and lower in the AI condition compared to the no-AI condition. No other main or interaction effects were significant.

5 | Discussion

In today's academic environment, the rapid diffusion of GenAI is reshaping not only how knowledge is produced, but also how those who produce it are perceived. Drawing on a sample of 651 participants, this study investigates how the use of GenAI—a technology that may be socially interpreted as masculine—differentially shapes perceptions of competence and warmth

across gender and across two task contexts (teaching and research) within academia.

Existing empirical studies have predominantly explored AI adoption behaviors enacted by women (Adewale 2025), the barriers they encounter (Buele and Llerena-Aguirre 2025), or the disadvantages generated by algorithms (Shrestha and Das 2022). What remains largely unaddressed is how the use of Generative AI shapes perceptions of competence and warmth, and how these perceptions vary by gender and task domain. Our study advances the debate by shifting the analytical lens from access and use to social evaluation, treating GenAI as a socially meaningful evaluative cue that is interpreted through gender stereotypes and task expectations, and whose adoption becomes part of the evaluative repertoire through which women's academic work is perceived. In adopting this lens, our study does not examine the situated construction or negotiation of these meanings; rather, it examines how such preexisting associations inform the evaluation of individuals who use GenAI.

Consistent with this SCM framing (Cuddy et al. 2008; Fiske et al. 2002; Fiske et al. 2007), we found that GenAI adoption is positively associated with perceived competence (H1a) and negatively associated with perceived warmth (H1b). This pattern supports the idea that technology use functions as a signal of analytical and innovative capability (Cheryan et al. 2015; Compagni et al. 2015), confirming that artifacts are not merely functional but carry socially shared associations with expertise that may inform evaluations of those who use them. Adewale (2025), for instance, shows that women themselves perceive GenAI as productivity-enhancing but ethically risky; our results extend this insight by showing that this ambivalence

is not only a feature of how women experience GenAI, but also of how they are perceived through it.

The two-way interactions between GenAI use and gender were not significant for either competence (H2a) or warmth (H2b): accordingly, neither hypothesis was supported at the level of the omnibus interaction test. Post hoc comparisons, however, revealed a pattern consistent with the theoretical expectations underlying H2b, although they did not provide formal support for the hypothesized interaction effect. Although AI adoption is associated with a decreased perceived warmth overall, this penalty was significant only for women both in teaching and research. This pattern is consistent with the gendered expectations of warmth, care, and relational availability identified in our argument: men, already seen as less warm, were not affected, whereas women who are held to stricter communal expectations, face warmth penalties when adopting technologies that signal masculine-coded competence (Rudman and Glick 2001).

Importantly, the three-way interaction between task, gender, and AI adoption on competence ratings suggests that the moderating role of gender is itself a function of the domain. In research, GenAI adoption appears to increase perceived competence for women but not for men, supporting H3a. In teaching, GenAI use is associated with greater perceived competence for both genders, although the pattern is stronger for men. In this sense, women appear to gain more competence in agentic tasks because GenAI offsets stereotypes of lower technical ability—a pattern that resonates with the “*prove it again bias*” whereby women must repeatedly demonstrate competence to be granted legitimacy (J. C. Williams et al. 2016). For men—whose technical competence in research is already presumed (Wajcman 2004)—GenAI adds a less evaluative value. Coherently, men who do not adopt GenAI are more penalized in teaching than in research, whereas no such asymmetry emerges for women. Taken together, gender moderates the competence effect of GenAI only when the alignment among gender stereotypes, task expectations, and the socially shared associations attached to technology use is considered simultaneously.

Turning to warmth, H3b was not supported at the level of the omnibus interaction: the three-way interaction between GenAI use, gender, and task was not significant. However, post hoc comparisons reveal a pattern consistent with the theoretical logic underlying H3b: women using GenAI face stronger warmth penalties in research than in teaching, and even without GenAI, women are perceived as significantly less warm in research than in teaching. Although these comparisons do not formally support H3b, they suggest a pattern of double punishment: women lose warmth both for engaging in research, a masculine-coded domain, and for adopting GenAI within it. This pattern is consistent with the persistent association between research and masculinity (Galvin et al. 2024). Women's presence in research contexts, regardless of technology use, may already be perceived as less congruent with communal expectations and be associated with lower warmth evaluations. When GenAI is added, the post hoc pattern suggests that these penalties may compound, as women cross multiple gender boundaries simultaneously, working in a masculine domain while using masculine technology. This pattern resonates with the

framing of academia as a context in which evaluations are not neutral (Acker 1990; van den Brink and Benschop 2012), and in which research is not a neutral evaluative space but a regime in which credibility, authority, and professional legitimacy are unevenly distributed (van Helden et al. 2023), and where the perceptual cost of using a masculine-coded tool may concentrate precisely where institutional rewards are highest.

Taken alongside prior work, Adewale (2025) and Buele and Llerena-Aguirre (2025) suggest that AI tools may ease women's research workload and partially mitigate structural disadvantages, whereas Ali et al. (2026) argue that algorithmic systems mostly reproduce existing inequalities. Our results identify an additional, perceptual channel through which inequality is reproduced—one that these studies do not isolate: women who use AI to enter or thrive in masculine-coded domains may receive the competence dividend the literature predicts, but this competence dividend may be accompanied by a social cost in warmth, a cost that is rarely measured and seldom theorized.

These findings directly address the central question motivating this study: whether emerging technologies reproduce, reshape, or intensify gendered evaluations. They suggest that GenAI may do all three simultaneously, but along different evaluative dimensions. They indicate a more cautious reading of recent claims that AI tools may help close gendered gaps in research productivity (Adewale 2025). The competence boost is real but partial, and the post hoc comparisons suggest that it may be accompanied by a warmth cost, particularly in the domain where the productivity gains are most needed. This extends the double bind framework (Jamieson 1995) by suggesting that women in knowledge work struggle to balance competence and warmth, with tradeoffs particularly severe in male-dominated domains and when masculine-coded technological tools are employed.

5.1 | Theoretical Implications

Our study contributes theoretically in three ways. First, it contributes to innovation literature by specifying the social-evaluative role of GenAI in gendered evaluation and advances the view of technology as a socially determined artifact. Technologies are not neutral instruments but social artifacts that incorporate and elicit specific meanings, values, and beliefs (Faulkner 2001; Leonardi 2011; Orlikowski 1999; Wajcman 2004), and these socially shared meanings may inform the evaluation of those who adopt them. As a technology stereotypically associated with rationality, technical control, and innovation, GenAI may be socially interpreted as carrying values that have long been culturally coded as masculine (Cheryan et al. 2015; Lohan and Wendy 2004; Wajcman 1991, 2004); when women adopt it, these associations are interpreted in relation to gender expectations in ways that shape how their competence and warmth are read. Our findings suggest that GenAI adoption may simultaneously operate as a competence-enhancing signal and as a warmth-penalizing signal, with the post hoc patterns suggesting that the warmth penalty may vary by gender and academic domain. The same artifact, in other words, may be perceived differently depending on the gender of its user, and the direction and intensity of these perceptions appear to depend on the alignment among the socially shared

meaning of the technology, the socially shared meaning of the task, and the gender of the user.

This contribution also shifts the analytical focus of feminist technology studies. Work in this tradition has long understood technologies as cultural artifacts that encode and reproduce gendered values, hierarchies, and worldviews. Our findings suggest that, even when algorithmic outputs are not at issue, the social meaning attached to GenAI use can itself inform gendered evaluations, operating through the perception of users rather than through the behavior of systems. In this sense, inequality may be reproduced not only through biased systems, but also through the ways in which users of these systems are perceived. By doing so, we shift attention from the technical black box of algorithmic functioning to the social black box of evaluation: the process through which socially shared associations attached to GenAI use inform perceptions of users' competence, warmth, and legitimacy. Hence, AI does not enter academic evaluation as a neutral instrument; rather, it may amplify the gendered logics already embedded in academia, understood as a gendered organizational field (Acker 1990; O'Connor 2020; van den Brink and Benschop 2012), where the evaluative regimes of credibility, authority, and legitimacy are already unevenly distributed (Guarino and Borden 2017; Herschberg et al. 2018; Järvinen and Mik-Meyer 2025a, 2025b).

Second, our findings have implications for the SCT and RCT. The SCM understands competence and warmth as relatively stable dimensions through which individuals and groups are socially evaluated (Cuddy et al. 2008; Fiske et al. 2002), and recent extensions of the model have positioned technological artifacts mainly as objects of evaluation rather than as forces that shape it (Kervyn et al. 2010; McKee et al. 2023). Our findings suggest that technology may also operate as a socially meaningful evaluative cue that shapes how users are evaluated along these same dimensions. GenAI does not appear to enter evaluation simply as another target to be rated along the competence and warmth dimensions; rather, it may operate as a socially meaningful signal that informs how the user is evaluated along those same dimensions. For women, this process appears to shift competence and warmth in opposite directions simultaneously, whereas for men it appears to move competence primarily. This suggests that SCM can be further extended by considering how technological artifacts condition evaluations differently depending on the gender of the user and the socially shared meaning of the task domain. In this sense, our findings point to the importance of recognizing the fact that stereotype activation is conditional on the alignment among user, task, and artifact coding, and that the warmth cost of competence repair may become more pronounced with the socially shared masculine coding of the field in which adoption takes place. Thus, the contribution is not only to show that GenAI may be evaluated as a cold-and-competent artifact, but also to specify how such artifact-based associations may inform the evaluation of those who use it (Aue et al. 2021; Bigman and Gray 2018). Our findings also refine RCT: it understands the evaluative disadvantage women face in agentic roles as a violation of communal expectations (Eagly and Karau 2002; Rudman 1998; Rudman and Glick 2001), and this prediction has typically been investigated in leadership and selection settings, where the violation is enacted through behavior. Our findings suggest that

this violation may also be triggered by the artifacts a woman adopts: in our study, the use of a masculine-coded technology was associated with post hoc patterns consistent with role-incongruity penalties on the warmth dimension, even when the task itself remains the same. At the same time, our findings invite a more careful engagement with a tension that recent optimistic accounts of GenAI as an equalizer for women in academia have not fully addressed (Adewale 2025; Buele and Llerena-Aguirre 2025). From an RCT perspective, the competence gains associated with GenAI may simultaneously activate communal violations that generate warmth penalties. Our results are consistent with this reading and add a specification the theory had not previously articulated: role incongruity does not resolve through technology adoption but may be redistributed across the competence and warmth dimensions, and the post hoc patterns suggest that the warmth side of this redistribution may concentrate in the field where the competence gain is largest. In this sense, the double bind described by Jamieson (1995) does not dissolve when women adopt new tools; it migrates with them.

Third, our study contributes to understanding academia as an organizational field in which evaluative regimes structure the recognition of competence, warmth, and legitimacy. Drawing on prior literature, universities are not gender-neutral meritocratic arenas but gendered organizations in which apparently neutral criteria of value, commitment, and authority reproduce gendered hierarchies (Acker 1990; van den Brink and Benschop 2012; O'Connor 2020). Within this regime, academic work is internally differentiated: research is associated with excellence, originality, and intellectual authority, whereas teaching and service are associated with availability, care, and relational labor—domains that are neither equivalent in prestige nor gender-neutral (Guarino and Borden 2017; Herschberg et al. 2018; Järvinen and Mik-Meyer 2025a, 2025b). Our findings show that GenAI use enters this regime as an additional evaluative signal that is interpreted through, and may amplify, these preexisting logics. The post hoc pattern of a compounded warmth penalty for women in research, alongside the three-way interaction on competence, suggests that the evaluative consequences associated with GenAI use may vary by domain, and that such warmth penalties may concentrate where institutional rewards are highest. In this sense, academia is not a neutral backdrop in which gendered evaluations of GenAI happen to occur, but an analytically generative context in which social-evaluative, stereotypical, and institutional mechanisms unfold together to produce gendered evaluative outcomes.

5.2 | Practical Implications

Beyond its theoretical contribution, this study provides practical implications for academic leaders seeking to address the gendered dynamics of GenAI use. Our study shows that GenAI adoption produces both opportunities and penalties, with the significant competence interaction and the post hoc warmth patterns indicating that these consequences may vary across gender and task domains, and that these competence and warmth dynamics call for distinct recommendations.

Decades of research show that many of the most common organizational interventions, such as diversity training,

awareness campaigns, and generic mentoring programs, produce limited and sometimes counterproductive effects (Dobbin and Kalev 2013, 2016; Calluso and Devetag 2026). More recently, McEwen et al. (2026) have argued that diversity, equality, and inclusion initiatives risk becoming “*repressive equality regimes*” that are formally inclusive but functionally depoliticizing, using the language of equality without disturbing the structures that produce inequality (Calluso and Iacopino 2026). Against this backdrop, our findings support two specific recommendations for academic leaders.

The first recommendation is related to “*where*” intervention should be addressed. Our results indicate that the gendered evaluative consequences of GenAI use are produced at the moment of evaluation, not at the moment of adoption: it is evaluators who interpret GenAI use through gendered schemas, not women’s decisions to adopt or disclose that generate the evaluative patterns identified in our analyses. This locates the appropriate site of intervention in concrete evaluative practices—hiring committees, promotion reviews, peer evaluation, and student assessment—rather than in policies that target users. By contrast, AI-use disclosure rules, AI literacy campaigns for women, or interventions aimed at encouraging women’s AI adoption are currently among the most visible policy responses, reflecting a literature focused on closing gender gaps in AI familiarity and use. Gesser-Edelsburg et al. (2024), for instance, find that men report significantly higher familiarity with GenAI tools than women (27% vs. 13%), even though actual usage rates do not differ; accordingly, the authors recommend AI literacy initiatives, mentorship programs, and curriculum integration to bridge this divide. While such interventions address an important dimension of the gender–AI relationship, they do not address the perceptual channel we identify and may even reinforce it by making GenAI use a more visible evaluative cue.

The second recommendation is related to “*how*” intervention should be designed. Existing research suggests that interventions which constrain evaluators’ discretion and create accountability for evaluative outcomes are more effective than those that seek to change evaluators’ attitudes (Castilla 2008; Dobbin and Kalev 2016). Applied to our work, this points to evaluative architectures that make GenAI use a less salient cue and competence judgments more directly tied to substantive work, such as prespecified, criterion-based rubrics in hiring and promotion; blinded or partially blinded review where feasible; calibration sessions in which evaluators compare judgments and surface inconsistencies; and the routine auditing of evaluative outcomes disaggregated by gender and task domain. Together, these practices reduce the space in which gendered interpretations of GenAI use can shape evaluative outcomes.

5.3 | Limitation and Future Research

Although this study provides novel insights into the gendered perceptions associated with GenAI use in academia, several limitations point to important avenues for future research.

First, the study conceptualizes GenAI primarily as a socially meaningful cue of technological adoption, without differentiating between specific applications or uses. This focus reflects

the study’s aim to open the social-evaluative, rather than the technical, black box of GenAI by examining how observable adoption shapes professional judgments. At the same time, different forms of GenAI use—such as writing assistance, data analysis, or coding support—may carry distinct professional meanings, normative expectations, and ethical implications within academic work, and may therefore be evaluated differently by academic audiences (Andersen et al. 2025). Distinguishing among these applications would instead address the technical black box and whether specific uses generate different evaluations. Future research could further unpack these distinctions to examine how different uses of GenAI are differently interpreted and evaluated across academic domains, thereby complementing our analysis of the social-evaluative process through which GenAI adoption itself becomes a meaningful cue in professional evaluation.

Second, the vignette-based experimental design enables the identification of causal relationships through controlled manipulation but necessarily reduces the complexity of real-world contexts in order to isolate the variables of interest (Skilling and Stylianides 2020). As such, our study captures evaluative perceptions rather than enacted practices. This distinction is particularly relevant in academic settings, where the use of GenAI is often negotiated, contested, and embedded in local norms. Future research could therefore adopt complementary methodological approaches—such as qualitative or mixed-method designs—to explore how these perceptions translate into situated practices. Furthermore, our study offers a valuable cross-sectional snapshot of the gendered perceptions surrounding GenAI in academia, but this emerging technology and the discursive frames through which it is understood are evolving rapidly. Perception captured at this stage may shift as GenAI use become routinized. Replication and longitudinal extensions of this study will be important for understanding whether the gendered perceptions we identify are fixed features of academic culture or artifacts of an early-stage technological transition.

Third, the socially shared meanings associated with GenAI use as an evaluative cue are likely to be shaped by broader cultural and institutional contexts. Although the sample includes participants from multiple countries, it is predominantly drawn from Western academic systems, in which dominant technological imaginaries and gendered expectations are deeply intertwined (Čičigoj et al. 2025). Given that both technological imaginaries and gendered expectations are culturally contingent (Calluso et al. 2026), the generalizability of the findings beyond these contexts remains an open question. Future research should therefore examine how these dynamics unfold across different cultural and institutional settings. Furthermore, perceptions of GenAI use also intersect with users’ social positions beyond gender, such as disability, race, career stage, and disciplinary affiliation, and recent studies suggest that AI technologies are entangled with structures of power and exclusion that shape both who is seen as a legitimate user and how their use is evaluated (Foley and Melese 2025). Future research should examine how these dynamics unfold in non-Western systems and across other social dimensions that interact with gender.

Finally, although teaching and research were operationalized as proxies for communal and agentic task domains, these

categories inevitably simplify the heterogeneity of academic work. Academic labor encompasses a wide range of activities that may not neatly align with these distinctions. Future studies could develop more fine-grained operationalizations to better capture how different dimensions of academic work intersect with gendered evaluations and the socially shared meanings attached to technology use.

6 | Conclusions

This study shows that Generative AI (GenAI) adoption in academia has gendered effects on perceptions of competence and warmth. GenAI use enhances perceived competence for all academics, but women gain most in agentic research tasks, whereas men's evaluations are less affected due to existing stereotypes. Conversely, GenAI adoption reduces perceived warmth, with this penalty appearing to affect women more strongly and to be particularly pronounced in research. This pattern is consistent with a "double bind" in which women's competence gains may come at the expense of relational perceptions.

In conclusion, although GenAI offers opportunities to signal expertise, it may also reinforce gendered tradeoffs, underscoring the need for institutional and cultural interventions to ensure technological proficiency benefits all academics equitably.

Acknowledgments

The authors are deeply grateful to the guest editors of the special issue—Alessandra Lazazzara (University of Milano, Italy), Efraxia D. Zamani (Durham University Business School, UK), Andri Georgiadou (University of Nottingham, UK), and Ayushi Tandon (Trinity College Dublin, Ireland)—for their thoughtful guidance and valuable feedback on earlier versions of this work. The authors used ChatGPT (OpenAI 2025), a generative AI-based large language model, exclusively for proofreading and language refinement. The tool was not used in the conceptualization, analysis, interpretation, or development of the study's findings. The authors reviewed and approved all revisions and take full responsibility for the final content of the manuscript. Open access publishing facilitated by Libera Università Internazionale degli Studi Sociali Guido Carli, as part of the Wiley - CRUI-CARE agreement.

Funding

The authors have nothing to report.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are openly available in Mendeley Data at <https://doi.org/10.17632/269kzdtj5s.1>.

References

- Acker, J. 1990. "Hierarchies, Jobs, Bodies: A Theory of Gendered Organizations." *Gender & Society* 4, no. 2: 139–158. <https://doi.org/10.1177/089124390004002002>.
- Acker, J. 1998. "Hierarchies, Jobs, Bodies: A Theory of Gendered Organizations." In *Feminist Foundations: Toward Transforming Sociology*, edited by K. A. Myers, C. D. Anderson, and B. J. Risan, 299–317. Sage.

- Adewale, S. 2025. "Exploring ChatGPT Usage Amongst Female Academics and Researchers in the Academia." *International Journal of Information and Learning Technology* 42, no. 1: 49–66. <https://doi.org/10.1108/IJILT-01-2024-0012>.
- Aguinis, H., and K. J. Bradley. 2014. "Best Practice Recommendations for Designing and Implementing Experimental Vignette Methodology Studies." *Organizational Research Methods* 17, no. 4: 351–371. <https://doi.org/10.1177/1094428114547952>.
- Ali, A., M. Aatif, and A. Haider. 2026. "Procedural Justice in the Era of Algorithms: A Critical Evaluation of AI Governance and Digital Rights." *International Journal of Social Sciences Bulletin* 4, no. 1: 1408–1424. <https://doi.org/10.5281/zenodo.18586156>.
- Andersen, J. P., L. Degn, R. Fishberg, et al. 2025. "Generative Artificial Intelligence (GenAI) in the Research Process—A Survey of Researchers' Practices and Perceptions." *Technology in Society* 81, no. June: 102813. <https://doi.org/10.1016/j.techsoc.2025.102813>.
- Araneda-Guirriman, C., G. Sepúlveda-Páez, L. Pedraja-Rejas, and J. S. Martín. 2023. "Women in Academia: An Analysis Through a Scoping Review." *Frontiers in Education* 8: 1137866. <https://doi.org/10.3389/edu.2023.1137866>.
- Atzmüller, C., and P. M. Steiner. 2010. "Experimental Vignette Studies in Survey Research." *Methodology* 6, no. 3: 128–138. <https://doi.org/10.1027/1614-2241/a000014>.
- Aue, T., S. Bühner, B. Mayer, and M. Dricu. 2021. "Empathic Responses to Social Targets: The Influence of Warmth and Competence Perceptions, Situational Valence, and Social Identification." *PLoS One* 16, no. 3: e0248562. <https://doi.org/10.1371/journal.pone.0248562>.
- Barocas, S., and A. D. Selbst. 2016. "Big Data's Disparate Impact." *California Law Review* 104: 671–732. <https://doi.org/10.15779/Z38BG31>.
- Barry, I., and E. Stephenson. 2025. "The Gendered, Epistemic Injustices of Generative AI." *Australian Feminist Studies* 40, no. 123: 1–21. <https://doi.org/10.1080/08164649.2025.2480927>.
- Benschop, Y., and M. Brouns. 2003. "Crumbling Ivory Towers: Academic Organizing and Its Gender Effects." *Gender, Work and Organization* 10, no. 2: 194–212. <https://doi.org/10.1111/1468-0432.t01-i-00011>.
- Berger, J., C. L. Ridgeway, M. Hamit Fisek, and R. Z. Norman. 1998. "The Legitimation and Delegitimation of Power and Prestige Orders." *American Sociological Review* 63, no. 3: 379–405. <https://doi.org/10.2307/2657555>.
- Biernat, M., and K. Fiegen. 2001. "Shifting Standards and the Evaluation of Competence: Complexity in Gender-Based Judgment and Decision Making." *Journal of Social Issues* 57, no. 4: 707–724. <https://doi.org/10.1111/0022-4537.00237>.
- Bigman, Y. E., and K. Gray. 2018. "People Are Averse to Machines Making Moral Decisions." *Cognition* 181, no. (December): 21–34. <https://doi.org/10.1016/j.cognition.2018.08.003>.
- Blalock, A. E., and A. Stefanese-Yates. 2024. "Hierarchies and Paradoxes: How Women in non-tenure-track Faculty Positions Experience a Gendered Organization." *Journal of Diversity in Higher Education* 17, no. 6: 856–871. <https://doi.org/10.1037/dhe0000486>.
- Browne, J., S. Cave, E. Drage, and K. McInerney. 2023. *Feminist AI: Critical Perspectives on Algorithms, Data, and Intelligent Machines*. Oxford University Press. <https://doi.org/10.1093/oso/9780192889898.001.0001>.
- Buele, J., and L. Llerena-Aguirre. 2025. "Transformations in Academic Work and Faculty Perceptions of Artificial Intelligence in Higher Education." *Frontiers in Education* 10: 1603763. <https://doi.org/10.3389/edu.2025.1603763>.
- Calder, B. J., L. W. Phillips, and A. M. Tybout. 1981. "Designing Research for Application." *Journal of Consumer Research* 8, no. 2: 197–207. <https://doi.org/10.1086/208856>.

- Calluso, C., and G. Devetag. 2026. "When Equality Policies Backfire: Structural Authority and the Cognitive Effects of ESG Gender Governance." *SSRN working paper*. <https://doi.org/10.2139/ssrn.6385142>.
- Calluso, C., and V. Iacopino. 2026. "When Egalitarianism Legitimizes Inequality: Implications for Gender Disparities in Organizations." *Behavioral and Brain Sciences* 49: e424. <https://doi.org/10.1017/S0140525X26105871>.
- Calluso, C., V. Iacopino, and A. Piazza. 2026. "The Impact of Intercultural Differences and Gender on Respectful Leadership and Employee Engagement." *European Journal of International Management* 1, no. 1: 10079097. <https://doi.org/10.1504/EJIM.2026.10079097>.
- Castilla, E. J. 2008. "Gender, Race, and Meritocracy in Organizational Careers." *American Journal of Sociology* 113, no. 6: 1479–1526. <https://doi.org/10.1086/588738>.
- Cheryan, S., M. Allison, and A. N. Meltzoff. 2015. "Cultural Stereotypes as Gatekeepers: Increasing Girls' Interest in Computer Science and Engineering by Diversifying Stereotypes." *Frontiers in Psychology* 6: 49. <https://doi.org/10.3389/fpsyg.2015.00049>.
- Čičigoj, K., N. Majsova, and J. Šepetavc. 2025. "Peripheral Visions of Alternative Futures: Feminist Techno-Imaginations." *Feminist Encounters: A Journal of Critical Studies in Culture and Politics* 9, no. 2: 20. <https://doi.org/10.20897/femenc/16778>.
- Cockburn, C. 1988. "Machinery of Dominance: Women, Men and Technical Know-How." *Cahiers de l'APRE* 7, no. 1: 93–99. <https://doi.org/10.3406/apre.1988.864>.
- Cockburn, C. 1991. *In the Way of Women: Men's Resistance to Sex Equality in Organizations*. ILR Press.
- Cockburn, C. 1993. "Technical Competence, Gender Identity and Women's Autonomy." *Cahiers du GEDISST* 7: 111–122. <https://doi.org/10.3406/genre.1993.921>.
- Cockburn, C. 1997. "Domestic Technologies: Cinderella and the Engineers." *Women's Studies International Forum* 20, no. 3: 361–371. [https://doi.org/10.1016/S0277-5395\(97\)00020-4](https://doi.org/10.1016/S0277-5395(97)00020-4).
- Cockburn, C. 2009. "On 'The Machinery of Dominance: Women, Men, and Technical Know-How.'" *WSQ: Women's Studies Quarterly* 37, no. 1–2: 269–273. <https://doi.org/10.1353/wsqs.0.0148>.
- Cockburn, C. 2025. "The Gendering of Technology." In *Understanding Technology in Education*, edited by H. Mackay, M. Young, and J. Beynon, 41–65. Falmer Press.
- Compagni, A., V. Mele, and D. Ravasi. 2015. "How Early Implementations Influence Later Adoptions of Innovation: Social Positioning and Skill Reproduction in the Diffusion of Robotic Surgery." *Academy of Management Journal* 58, no. 1: 242–278. <https://doi.org/10.5465/amj.2011.1184>.
- Connor, R. A., and S. T. Fiske. 2017. "Warmth and Competence: A Feminist Look at Power and Negotiation." In *APA Handbook of the Psychology of Women: History, Theory, and Battlegrounds*, edited by C. B. Travis, J. W. White, A. Rutherford, W. S. Williams, S. L. Cook, and K. F. Wyche, Vol. 1, 321–342. American Psychological Association. <https://doi.org/10.1037/0000059-016>.
- Cuddy, A. J. C., S. T. Fiske, and P. Glick. 2008. "Warmth and Competence as Universal Dimensions of Social Perception: The Stereotype Content Model and the BIAS Map." *Advances in Experimental Social Psychology* 40: 61–149. [https://doi.org/10.1016/S0065-2601\(07\)00002-0](https://doi.org/10.1016/S0065-2601(07)00002-0).
- Cuddy, A. J. C., S. T. Fiske, V. S. Y. Kwan, et al. 2009. "Stereotype Content Model Across Cultures: Towards Universal Similarities and Some Differences." *British Journal of Social Psychology* 48, no. 1: 1–33. <https://doi.org/10.1348/014466608X314935>.
- D'Ignazio, C., and L. F. Klein. 2020. *Data Feminism*. MIT Press.
- Dobbin, F., and A. Kalev. 2013. "The Origins and Effects of Corporate Diversity Programs." In *The Oxford Handbook of Diversity and Work*, edited by Q. M. Roberson, 253–281. Oxford University Press.
- Dobbin, F., and A. Kalev. 2016. "Why Diversity Programs Fail." *Harvard Business Review* 94, no. 7–8: 52–60. <https://hbr.org/2016/07/why-diversity-programs-fail>.
- Dorta-González, P., A. Jorge López-Puig, M. I. Dorta-González, and S. M. González-Betancor. 2024. "Generative Artificial Intelligence Usage by Researchers at Work: Effects of Gender, Career Stage, Type of Workplace, and Perceived Barriers." *Telematics and Informatics* 94, no. (October): 102187. <https://doi.org/10.1016/j.tele.2024.102187>.
- Eagly, A. H., and S. J. Karau. 2002. "Role Congruity Theory of Prejudice Toward Female Leaders." *Psychological Review* 109, no. 3: 573–598. <https://doi.org/10.1037/0033-295X.109.3.573>.
- Faulkner, W. 2001. "The Technology Question in Feminism: A View From Feminist Technology Studies." *Women's Studies International Forum* 24, no. 1: 79–95. [https://doi.org/10.1016/S0277-5395\(00\)00166-7](https://doi.org/10.1016/S0277-5395(00)00166-7).
- Fiske, S. T., and A. J. C. Cuddy. 2005. "Stereotype Content Across Cultures as a Function of Group Status." In *Social Comparison and Social Psychology: Understanding Cognition, Intergroup Relations, and Culture*, edited by S. Guimond, 249–263. Cambridge University Press. <https://doi.org/10.1017/CBO9780511584329.013>.
- Fiske, S. T., A. J. C. Cuddy, and P. Glick. 2007. "Universal Dimensions of Social Cognition: Warmth and Competence." *Trends in Cognitive Sciences* 11, no. 2: 77–83. <https://doi.org/10.1016/j.tics.2006.11.005>.
- Fiske, S. T., A. J. C. Cuddy, P. Glick, and J. Xu. 2002. "A Model of (Often Mixed) Stereotype Content: Competence and Warmth Respectively Follow From Perceived Status and Competition." *Journal of Personality and Social Psychology* 82, no. 6: 878–902. <https://doi.org/10.1037/0022-3514.82.6.878>.
- Fiske, S. T., A. J. C. Cuddy, P. Glick, and J. Xu. 2018. "A Model of (Often Mixed) Stereotype Content: Competence and Warmth Respectively Follow From Perceived Status and Competition." In *Social Cognition: Selected Works of Susan Fiske*, 163–214. Routledge. <https://doi.org/10.4324/9781315187280-7>.
- Foley, A., and F. Melese. 2025. "Disabling AI: Power, Exclusion, and Disability." *British Journal of Sociology of Education*: 1–22. <https://doi.org/10.1080/01425692.2025.2519482>.
- Fortin, J., B. Bartlett, M. Kantar, M. Tseng, and Z. Mehrabi. 2021. "Digital Technology Helps Remove Gender Bias in Academia." *Scientometrics* 126, no. 5: 4073–4081. <https://doi.org/10.1007/s11192-021-03911-4>.
- Foschi, M., and S. Buchan. 1990. "Ethnicity, Gender, and Perceptions of Task Competence." *Canadian Journal of Sociology* 15, no. 1: 1–18. <https://doi.org/10.2307/3341170>.
- Galvin, D. J., S. C. Anderson, C. J. Marolf, N. G. Schneider, and A. L. Liebl. 2024. "Comparative Analysis of Gender Disparity in Academic Positions Based on U.S. Region and STEM Discipline." *PLoS One* 19, no. 3: e0298736. <https://doi.org/10.1371/journal.pone.0298736>.
- Gesser-Edelsburg, A., H. Rana, E. Eliyahu, and A. Tal. 2024. "Bridging the Divide: An Empirical Investigation of Artificial Intelligence and Generative Artificial Intelligence Integration Across Genders, Disciplines and Academic Roles." Supplement, *European Journal of Open, Distance and E-Learning* 26, no. S1: 51–69. <https://doi.org/10.2478/eurodl-2024-0008>.
- Gross, N. 2023. "What ChatGPT Tells Us About Gender: A Cautionary Tale About Performativity and Gender Biases in AI." *Social Sciences* 12, no. 8: 435. <https://doi.org/10.3390/socsci12080435>.
- Guarino, C. M., and V. M. H. Borden. 2017. "Faculty Service Loads and Gender: Are Women Taking Care of the Academic Family?" *Research in Higher Education* 58, no. 6: 672–694. <https://doi.org/10.1007/s11162-017-9454-2>.

- Haraway, D. 1991. "A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late Twentieth Century." In *Simians, Cyborgs, and Women: The Reinvention of Nature*, 149–181. Routledge.
- Heijstra, T. M., F. S. Steinþórsdóttir, and P. J. Einarsdóttir. 2017. "Academic Career Making and the Double-Edged Role of Academic Housework." *Gender and Education* 29, no. 6: 764–780. <https://doi.org/10.1080/09540253.2016.1171825>.
- Herschberg, C., Y. Benschop, and M. van den Brink. 2018. "Selecting Early-Career Researchers: The Influence of Discourses of Internationalisation and Excellence on Formal and Applied Selection Criteria in Academia." *Higher Education* 76, no. 5: 807–825. <https://doi.org/10.1007/s10734-018-0237-2>.
- Highhouse, S. 2009. "Designing Experiments That Generalize." *Organizational Research Methods* 12, no. 3: 554–566. <https://doi.org/10.1177/1094428107300396>.
- Hsiao, C. H., and K. Yu Tang. 2025. "Beyond Acceptance: An Empirical Investigation of Technological, Ethical, Social, and Individual Determinants of GenAI-Supported Learning in Higher Education." *Education and Information Technologies* 30, no. 8: 10725–10750. <https://doi.org/10.1007/s10639-024-13263-0>.
- Jamieson, K. H. 1995. *Beyond the Double Bind: Women and Leadership*. Oxford University Press.
- Järvinen, M., and N. Mik-Meyer. 2025a. "At the Heart or on the Periphery: Gender, (In)Visibility and Epistemic Positioning in Academia." *Gender and Education* 37, no. 4: 437–453. <https://doi.org/10.1080/09540253.2024.2442923>.
- Järvinen, M., and N. Mik-Meyer. 2025b. "Giving and Receiving: Gendered Service Work in Academia." *Current Sociology* 73, no. 3: 302–320. <https://doi.org/10.1177/00113921231224754>.
- Kalluri, P. 2020. "Don't Ask if Artificial Intelligence Is Good or Fair, Ask How it Shifts Power." *Nature* 583, no. 7815: 169. <https://doi.org/10.1038/d41586-020-02003-2>.
- Kervyn, N., Y. Vincent, and C. M. Judd. 2010. "Compensation Between Warmth and Competence: Antecedents and Consequences of a Negative Relation Between the Two Fundamental Dimensions of Social Perception." *European Review of Social Psychology* 21, no. 1: 155–187. <https://doi.org/10.1080/13546805.2010.517997>.
- Krings, F., S. Sczesny, and A. Kluge. 2011. "Stereotypical Inferences as Mediators of Age Discrimination: The Role of Competence and Warmth." *British Journal of Management* 22, no. 2: 187–201. <https://doi.org/10.1111/j.1467-8551.2010.00721.x>.
- Lagesen, V. A., and I. Suboticki. 2024. "Performing Excellence and Gender Balance in Higher Education." *Higher Education* 88, no. 2: 683–701. <https://doi.org/10.1007/s10734-023-01138-1>.
- Leonardi, P. M. 2011. "Innovation Blindness: Culture, Frames, and Cross-Boundary Problem Construction in the Development of New Technology Concepts." *Organization Science* 22, no. 2: 347–369. <https://doi.org/10.1287/orsc.1100.0529>.
- Lin, W.-Y., J.-Wu Wang, H.-Yu Lin, H.-T. Lin, and B. T. Johnson. 2011. "When Low-Warmth Targets Are Liked: The Roles of Competence, Gender, and Relative Status." *Journal of Psychology: Interdisciplinary and Applied* 145, no. 3: 247–265. <https://doi.org/10.1080/00223980.2011.558132>.
- Lohan, M., and F. Wendy. 2004. "Masculinities and Technologies: Some Introductory Remarks." *Men and Masculinities* 6, no. 4: 319–329. <https://doi.org/10.1177/1097184X03260956>.
- McEwen, C., A. Pullen, and C. Rhodes. 2026. "Repressive Equality Regimes: How Diversity, Equality and Inclusion Initiatives Constrain Equality and Conceal Politics at Work." *Work, Employment & Society*. <https://doi.org/10.1177/0950017025141277>.
- McKee, K. R., X. Bai, and S. T. Fiske. 2023. "Humans Perceive Warmth and Competence in Artificial Intelligence." *iScience* 26, no. 8: 107256. <https://doi.org/10.1016/j.isci.2023.107256>.
- Nedungadi, P., M. Ramesh, V. Govindaraju, B. Rao, P. Berbeglia, and R. Raman. 2024. "Emerging Leaders or Persistent Gaps? Generative AI Research May Foster Women in STEM." *International Journal of Information Management* 77, no. (August): 102785. <https://doi.org/10.1016/j.ijinfomgt.2024.102785>.
- Nivenitha, D., and P. S. Nagarajan. 2025. "AI and Her: Bridging the Gender Gap Through Technology." In *Empowering Women: Bridging the Gap Through Entrepreneurship*, edited by S. Sudhamathy, 251–259. Alagappa Institute of Management, Alagappa University.
- Noble, S. U. 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York University Press.
- O'Connor, P. 2020. "Why Is it so Difficult to Reduce Gender Inequality in Male-Dominated Higher Educational Organizations? A Feminist Institutional Perspective." *Interdisciplinary Science Reviews* 45, no. 2: 207–228. <https://doi.org/10.1080/03080188.2020.1737903>.
- OpenAI. 2025. "Introducing GPT-5." August. <https://openai.com/index/introducing-gpt-5/>.
- Orlikowski, W. J. 1999. "Managing Use, Not Technology: A View From the Trenches." *Financial Times* 10–11. https://ocw.mit.edu/courses/15-343-managing-transformations-in-work-organizations-and-society-spring-2002/494b34de507aa7461598c3822d0fcc2_orlikowskimanagingusenottechnology.pdf.
- Orlikowski, W. J., and D. C. Gash. 1994. "Technological Frames: Making Sense of Information Technology in Organizations." *ACM Transactions on Information Systems* 12, no. 2: 174–207. <https://doi.org/10.1145/196734.196745>.
- Otis, N. G., S. Delecourt, K. Cranney, and R. Koning. 2024. "Global Evidence on Gender Gaps and Generative AI." *Harvard Business School Working Paper No. 25-023*. Harvard Business School.
- Phillips, M. J., P. L. Dzidic, and E. L. Castell. 2022. "Exploring and Critiquing Women's Academic Identity in Higher Education: A Narrative Review." *Sage Open* 12, no. 2: 21582440221096145. <https://doi.org/10.1177/21582440221096145>.
- Podolny, J. M. 2005. *Status Signals: A Sociological Study of Market Competition*. Princeton University Press. <https://doi.org/10.1515/978140837878>.
- Podsakoff, P. M., S. B. MacKenzie, J.-Y. Lee, and N. P. Podsakoff. 2003. "Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies." *Journal of Applied Psychology* 88, no. 5: 879–903. <https://doi.org/10.1037/0021-9010.88.5.879>.
- Rudman, L. A. 1998. "Self-Promotion as a Risk Factor for Women: The Costs and Benefits of Counterstereotypical Impression Management." *Journal of Personality and Social Psychology* 74, no. 3: 629–645. <https://doi.org/10.1037/0022-3514.74.3.629>.
- Rudman, L. A., and P. Glick. 2001. "Prescriptive Gender Stereotypes and Backlash Toward Agentic Women." *Journal of Social Issues* 57, no. 4: 743–762. <https://doi.org/10.1111/0022-4537.00239>.
- Shadish, W. R., T. D. Cook, and D. T. Campbell. 2002. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Houghton Mifflin.
- Shrestha, S., and S. Das. 2022. "Exploring Gender Biases in ML and AI Academic Research Through Systematic Literature Review." *Frontiers in Artificial Intelligence* 5: 976838. <https://doi.org/10.3389/frai.2022.976838>.
- Skilling, K., and G. J. Stylianides. 2020. "Using Vignettes in Educational Research: A Framework for Vignette Construction." *International Journal of Research and Method in Education* 43, no. 5: 541–556. <https://doi.org/10.1080/1743727X.2019.1704243>.

Storey, V. C., W. T. Yue, J. Leon Zhao, and R. Lukyanenko. 2025. "Generative Artificial Intelligence: Evolving Technology, Growing Societal Impact, and Opportunities for Information Systems Research." *Information Systems Frontiers* 27, no. 5: 2081–2102. <https://doi.org/10.1007/s10796-025-10581-7>.

Tang, C., S. Li, S. Hu, F. Zeng, and Q. Du. 2025. "Gender Disparities in the Impact of Generative Artificial Intelligence: Evidence From Academia." *PNAS Nexus* 4, no. 2: pgae591. <https://doi.org/10.1093/pnas/nexus/pgae591>.

Toupin, S. 2024. "Shaping Feminist Artificial Intelligence." *New Media & Society* 26, no. 1: 580–595. <https://doi.org/10.1177/14614448221150776>.

van den Brink, M., and Y. Benschop. 2012. "Gender Practices in the Construction of Academic Excellence: Sheep With Five Legs." *Organization* 19, no. 4: 507–524. <https://doi.org/10.1177/1350508411414293>.

van Helden, D. L., L. den Dulk, B. Steijn, and M. W. Vernooij. 2023. "Gender, Networks and Academic Leadership: A Systematic Review." *Educational Management Administration & Leadership* 51, no. 5: 1049–1066. <https://doi.org/10.1177/17411432211034172>.

Wajcman, J. 1991. *Feminism Confronts Technology*. Pennsylvania State University Press.

Wajcman, J. 2004. *TechnoFeminism*. Polity Press.

Wajcman, J., and E. Young. 2023. "Feminism Confronts AI: The Gender Relations of Digitalisation." In *Feminist AI: Critical Perspectives on Algorithms, Data, and Intelligent Machines*, edited by J. Browne, S. Cave, E. Drage, and K. McInerney, 47–64. Oxford University Press. <https://doi.org/10.1093/oso/9780192889898.003.0004>.

West, S. M., M. Whittaker, and K. Crawford. 2019. *Discriminating Systems: Gender, Race, and Power in AI*. AI Now Institute. <https://ainowinstitute.org/wp-content/uploads/2023/04/discriminatingystems.pdf>.

Williams, J. C., W. P. Katherine, and E. V. Hall. 2016. "Tools for Change: Boosting the Retention of Women in the STEM Pipeline." *Journal of Research in Gender Studies* 6, no. 1: 11–75. <https://doi.org/10.22381/JRG56120161>.

Williams, J. C., R. M. Korn, and S. Mihaylo. 2020. "Beyond Implicit Bias: Litigating Race and Gender Employment Discrimination Using Data From the Workplace Experiences Survey." *Hastings Law Journal* 72, no. 1: 337–464. https://repository.uclawsf.edu/hastings_law_journal/vol72/issue1/7/.

Williams, L. J., and A. Hervé. 2010. "Fisher's Least Significant Difference (LSD) Test." In *Encyclopedia of Research Design*, edited by N. J. Salkind, 840–853. Sage. <https://doi.org/10.4135/9781412961288.n154>.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.

Supporting Information S1: gwao70275-sup-0001-suppl-data.docx.