Paidéia
2025, Vol. 35, e3531. https://doi.org/10.1590/1982-4327e3531
ISSN 1982-4327 (online version)

Thematic Dossier

Sexism in University Teaching: Differences in the Evaluations of Male and Female Professors

Hyalle Abreu Viana[1]
Khalil da Costa Silva[2]
Elza Maria Techio[2]
José Luis Álvaro[3]
Ana Raquel Rosas Torres[4]

Abstract: this research's objective was to investigate whether the evaluation of the professional performance of university professors (variable operationalized as perceived fairness of the grade received) is influenced by the independent variables: (a) participant's gender, (b) professor's gender, (c) type of major that they teach, and (d) and the grade given to the students by professors. Two studies were conducted. Study 1 (N = 384), with a predictive-correlational design, and Study 2 (N = 376), with an experimental design. Both studies had university students as participants. Taken together, the results indicate that gender is related to perceived fairness of grades given by professors, and this effect varies depending on the type of professional context. These results point to gender inequality in university teaching and indicate that women are still perceived as different from men, especially in contexts considered "male."

Keywords: sexism, employee performance appraisal, employment discrimination

Sexismo na Docência Universitária: Diferenças nas Avaliações de Professores e Professoras

Resumo: O objetivo desta pesquisa foi investigar se a avaliação do desempenho profissional de professores universitários (variável operacionalizada como a justiça percebida da nota recebida) é influenciada pelas variáveis independentes: (a) gênero do participante, (b) gênero do professor, (c) tipo de curso em que lecionam e (d) nota atribuída pelos professores aos estudantes. Para tanto, foram realizados dois estudos, o Estudo 1 (N= 384), de desenho correlacional-preditivo, e o Estudo 2 (N= 376), com desenho experimental. Ambos tiveram estudantes universitários como participantes. Em conjunto, os resultados indicam que o gênero está associado à percepção de justiça das notas atribuídas por docentes e que essa relação varia a depender do tipo de contexto profissional. Esses resultados apontam a desigualdade de gênero na docência universitária e indicam que mulheres ainda são percebidas como diferentes dos homens, sobretudo em contextos considerados "masculinos".

Palavras-chave: sexismo, avaliação de desempenho profissional, discriminação no trabalho

Sexismo en la Docencia Universitaria: Diferencias en las Evaluaciones de Profesores y Professoras

Resumen: El objetivo de esta investigación fue investigar si la evaluación del desempeño profesional de los profesores universitarios (variable operacionalizada como la justicia percibida de la calificación recibida) está influenciada por las variables independientes: (a) género del participante, (b) género del profesor, (c) tipo de carrera en la que imparten clases y (d) calificación otorgada por los profesores a los estudiantes. Se llevaron a cabo dos estudios. El Estudio 1 (N = 384), de diseño correlacional-predictivo, y el Estudio 2 (N = 376), con diseño experimental. Ambos tuvieron estudiantes universitarios como participantes. En conjunto, los resultados indican que el género está relacionado con la percepción de equidad en las calificaciones asignadas por los profesores, y este efecto varía según el tipo de contexto profesional. Estos resultados apuntan a la desigualdad de género en la enseñanza universitaria e indican que todavía se perciben las mujeres como diferentes de los hombres, especialmente en contextos considerados "masculinos".

Palabras clave: sexismo, evaluación del rendimiento de empleados, discriminación en el trabajo

[1]*Centro Universitário Maurício de Nassau, João Pessoa - PB, Brazil.*
[2]*Universidade Federal da Bahia, Salvador-BA, Brazil.*
[3]*Universidad Complutense de Madrid, Madrid, Spain.*
[4]*Universidade Federal da Paraíba, João Pessoa - PB, Brazil.*

This article is the result of the first author's doctoral thesis, under the supervision of the fifth author. This research was approved by the Ethics Committee of the Universidade Federal da Paraíba (Brazil) under protocol no. 036/15. CAAE: 410228.0.0000.5188.

The study received financial support through a doctoral scholarship from CAPES.

Correspondence address: Khalil da Costa Silva. Instituto de Psicologia e Serviço Social. Rua Aristides Novis (Estrada de São Lázaro), 197, Federação. CEP 40210-730, Salvador – BA. E-mail: khalilsilva@ufba.br.

Members of social minorities, such as women, experience prejudice and discrimination on a daily basis in various areas of their lives (Hanek & Garcia, 2022; Heilman et al., 2024). Although progress has been made in some spaces and rights achieved, gender equality is still a long way off in our society (Hideg & Krstic, 2021). Regarding gender inequality in the workplace, one of the most blatant and common forms of discrimination against women is the marked disparity in the wage distribution between the sexes. Although women are the majority in higher education, their salaries are still lower than men's. In the formal sector, they receive 30% less than men (Instituto Brasileiro de Geografia e Estatística [IBGE], 2022). The 2024 Global Gender Gap Report found that it will take 152 years for wages, labor force participation, and opportunities for advancement to be equal for men and women (World Economic Forum, 2024). Among the countries surveyed, Brazil ranked 70th for gender equality, even though Brazilian laws specifically obligate equal pay between men and women (Lei No. 14.611, 2023; Liazibra, 2024).

These aspects of women in the labor market indicate that these wage inequalities cannot be fully attributed to differences in professional performance but have other reasons, including social norms based on the gender division of roles, which lead people to think and act in ways that maintain the status quo (Connor & Fiske, 2019; Tabassum & Nayak, 2021). This study seeks to address the following questions: "In the academic environment, does the gender of professors impact the evaluations they are given?", "Does the type of professional context (e.g., type of major they teach) also impact these evaluations?", and "Does the evaluation that a professor gives a student also impact the student's evaluation of the professor?"

It is assumed that the ideology of gender roles remains strong enough to justify the responsibilities and positions that men and women should occupy in our society (United Nations Development Programme, 2023). Gender stereotypes, understood as socially shared expectations about individuals' behavior based on their sex, are not only descriptive but also prescriptive (Heilman, 2012), they reflect historical roles in which women were more often associated with nurturing and domestic responsibilities while men assumed employment and provider roles (Eagly & Wood, 2012), with the former emphasizing helping, caring, and communal behaviors, and the latter requiring assertive, independent, and agentic tendencies (Eagly et al., 2000). The characterization of women as communal but not agentic conflicts with the demands of male gender-typed jobs, promoting perceptions of incompetence and penalties for counter-stereotypical behavior. Thus, women and men who break with the social norm and try to occupy contexts incongruent with the gender stereotypes must be penalized in some way. (Bareket et al., 2021; Costa et al., 2023; González et al., 2019; Heilman et al., 2024; Nyul et al., 2025; Viana et al., 2020).

Expósito et al. (1998) emphasize that research aimed at explaining discrimination against women at work based on socially shared gender stereotypes and beliefs has found interesting results. Taken together, these studies reveal that, despite a significant increase in the number of women entering areas that were previously reserved only for men, they continue to "prefer" professional occupations more congruent with the female stereotype. On the other hand, these authors also point to a loss of prestige in professions that have been attracting a growing number of women, confirming the ideology that men were born to perform high-status activities that enable them to rise, but women are not.

Subsequent investigations confirmed the findings of Expósito et al. (1998). Despite changes in public discourse surrounding gender issues and the increased participation of women in the labor market, research consistently identifies persistent inequalities in the labor market, rooted in gender stereotypes. Women remain disproportionately represented in professions related to caregiving or those requiring socio-emotional skills, whereas men are more commonly found in leadership positions or roles that demand logical-mathematical reasoning (Cortina et al., 2021; Costa et al., 2023; Fiske et al., 2002; Glick & Fiske, 1996). This separation sustained by the genders, according to Picanço et al. (2021), justified discrimination and legitimized the conduct adopted by men and women, especially in the workplace. This legitimization contributes to a psychosocial process in which the gender-based division in the labor market is perpetuated.

Within the educational field, gender stereotypes contribute to the perception that careers related to Science, Technology, Engineering, and Mathematics (STEM) are more appropriate for men, thereby perpetuating their predominance in these fields (Laksono & Patriot, 2021). As observed in other professional domains, gender-based divisions likely influence assessments of professional competence. Palazzo and Gomes (2009) state that for several decades, the educational and school systems have used various strategies to improve and encourage teaching performance using different forms of professional evaluation for admission and career progression by employing mechanisms such as tests and examinations; evaluations by peers, superiors, and students; career plans; and merit payment. In the case of evaluations conducted by students, the grade assigned by the professor is one of the factors that influences how students assess the professor's professional competence. Thus, when professors assign low grades to students, they tend to receive lower performance evaluations from those students (Valentín-Martínez & Mayor-Ruiz, 2023). Nevertheless, few studies have sought to investigate possible differences related to the professor's gender. To remedy this gap, Basow (2000) and Basow et al. (2006) demonstrated that, among other things, when asked to choose their best and worst professors, male students tend to choose men, and female students choose women as their best instructors. Aiming to specifically investigate this aspect, MacNell et al. (2015) analyzed gender bias in the evaluation of higher education students using a 2 × 2 design, in which the instructor's gender (real gender and manipulated gender) of an online class was experimentally manipulated. The results showed that, when asked to evaluate the instructors, the evaluated students gave lower marks to those perceived to be female than to those perceived to be male, regardless of

the instructor's quality of teaching or actual gender. In half of the classifications, women received much lower grades.

The cited studies indicate that the instructor's gender significantly impacts and is considered a guiding element that induces differences in the evaluations of male and female processes, converging with investigations that highlight gender-based discrimination in the workplace (Bareket et al., 2021; Costa et al., 2023; González et al., 2019; Heilman et al., 2024; Nyul et al., 2025; Viana et al., 2020). Nevertheless, there remains an aspect that requires further elucidation: Is it necessary for a woman to display a counter-stereotypical attribute or behavior, or does her mere presence in a traditionally male-dominated environment suffice to elicit discrimination? Given that, the objective of our study was to investigate whether the evaluation of the professional performance of university professors (variable operationalized as perceived fairness of the grade received) is influenced by the independent variables: (a) participant's gender, (b) professor's gender, (c) type of major that they teach, and (d) and the grade given to the students by professors. To achieve the objectives proposed here, two studies were developed. Study 1 examined whether the evaluation of professional performance was related to the participant's gender, the type of major, and the professor's gender. Study 2 expanded upon these findings by categorizing the professor's performance into two levels: positive and negative, operationalized as feedback provided on an academic task completed by the student.

Given the persistence of the gender-based social division of labor (Bareket et al., 2021; Costa et al., 2023; González et al., 2019; Heilman et al., 2024; Nyul et al., 2025; Viana et al., 2020), the general hypothesis posits that women will receive more negative evaluations than men, particularly when their work context contradicts the social norms associated with their gender roles.

Study 1 – Perceived fairness of the received grades

This study aimed to verify the relationship among the following variables: type of major (male or female), professor's gender (male or female), and participant's gender (male or female) on the perceived fairness by the students of the grade they received. A pilot study was carried out to determine which majors are perceived as typically masculine or feminine.

Method

Pilot Study

Participants

The 100 university students who participated had a mean age of 21.7 years (*SD*= 35.3 years), of which 52 were women and 48 were men. The sampling was non-probabilistic for convenience. The inclusion criteria were: accepting to participate in the study voluntarily and being a university student at a public institution.

Instrument

The *questionnaire* began with presenting a list of 42 undergraduate majors, and the participants were asked to categorize these majors as "what society considers to be typically male or female majors." For this purpose, participants were instructed to categorize each major in one of two categories: (1) masculine or (2) feminine. Then sociodemographic questions were asked.

Procedure

Data Collection. The questionnaire was answered individually and applied collectively in the classroom, after authorization from the teacher and signing of the Free and Informed Consent Form by the participant.

Data Analysis. Descriptive analyses were carried out using the SPSS version 22 software to verify the courses most frequently considered masculine or feminine by the participants.

Ethical Considerations

This work was approved by the Research Ethics Committee of Universidade Federal da Paraíba (Brazil) under the number 036/15. CAAE: 410228.0.0000.5188

Results

Engineering was considered male by 98% of participants, and the humanities (e.g., psychology) and some health majors (e.g., nursing, nutrition, etc.) were considered female by 97% of participants. Table 1 displays how the participants have categorized the majors. This categorization of male and female majors was used in Study 1.

Study 1

Participants

This is a study with a 2 (type of major) × 2 (professor's gender) × 2 (participant's gender) design. A total of 384

Table 1

Classification of Majors Perceived as Masculine or Feminine

Male Majors	Female Majors
Civil Engineering (98%)	Nursing and Nutrition (97%)
Computer Science, Physics, and Mathematics (97%)	Occupational Therapy (95%)
Electrical Engineering (95%)	Education (94%)
Agronomy (94%)	Speech-Language Pathology, Psychology, and Social Work (93%)
Physical Education and Economics (92%)	Languages (92%)

university students participated in this research, with ages ranging from 18 to 60 ($M = 23.6$; $SD = 6.47$), of which 60.7% were female. Sample size was previously calculated using G*Power software, estimating a sample of 237 participants as adequate for an experimental design with 80% power to detect a moderate effect size between the variables ($f = .25$). However, given the difficulty in finding female students in typically male majors and, simultaneously, male students in typically female majors, we sought to gather the largest possible number of participants we sought to collect as many participants as possible, within the concession granted by the universities. Participants were chosen for convenience and recruited within the majors perceived as "male" and "female" in the Pilot Study. The inclusion criteria were being a university student from a public institution and agreeing to participate in the study.

Instruments

The instrument consisted of two sections. In the first, the participants were asked which subjects they had received "the best grade" in the previous semester. Next, they were asked, using a seven-point Likert scale, how much they considered this grade fair or unfair (1= not at all fair and 7 = very fair). The second section contained sociodemographic questions: major, age, and gender of the participant.

Procedure

Data Collection. The professor's gender who taught the chosen subject was collected from the coordinators, because we had the name and period of the subject. This strategy was used to avoid any suspicion regarding the purpose of the study. We did not use the "worst grade" to avoid ambiguous situations regarding the professor's performance, such as students wanting revenge for having received a bad grade. The questionnaires were answered individually in the classroom.

Data Analysis. Using the SPSS version 22 software, an Analysis of Variance (ANOVA) was performed with a 2 x 2 x 2 design, taking as Independent Variables the type of major (Typically Male x Typically Female) X Professor's gender (Male x Female) X Participant gender (Male x Female) and as Dependent variable how much they considered this grade fair or unfair.

Ethical Considerations

This work was approved by the Research Ethics Committee of Universidade Federal da Paraíba (Brazil under the number 036/15. CAAE: 410228.0.0000.5188

Results

Based on the estimated means, the 2 × 2 × 2 ANOVA shows that there was a main effect of the professor's gender on the perceived fairness of the grade, $F(1,322) = 8.802$, $p = .003$, $\eta^2 = .027$, such that, when a male professor taught this subject, the grade received was seen to be even fairer ($M = 6.71$; $SD = 0.58$) than when it was assigned by a female professor ($M = 6.38$; $SD = 0.97$). The type of major $F(1,322) = .199$, $p = .65$, $\eta^2 = .001$ and the participant's gender, $F(1,322) = 2.865$, $p = .09$, $\eta^2 < .009$ did not produce an effect on the perceived fairness of the grade received.

The two-way interaction between the type of major and the participant's gender was not significant, $F(1, 322) = 1.77$, $p = .18$, $\eta^2 = .006$ The planned ANOVA comparisons, nevertheless, found differences only in the means of the participants in the female majors, $t(173) = -3.207$, $p = .002$, $d = 0.24$ indicating that the female students of these majors perceived the grade received as being fairer ($M = 6.76$; $SD = 0.46$) than the male students ($M = 6.32$; $SD = 1.3$). In male majors, no differences were observed, $t(146) = 0.60$, $p = .549$, $d = 0.05$.

The ANOVA showed an effect on the two-way interaction between the type of major and the professor's gender, $F(1,322) = 6.89$, $p = .009$, $\eta^2 = .021$ In the planned comparisons, statistically significant differences were found between the means of the participants in majors considered female, $t(158.39) = 1.949$, $p = .05$, $d = 0.15$. These results demonstrate that, in these majors, when the subject is taught by a male professor, the grade assigned by the professor is perceived as being fairer ($M = 6.75$; $SD = 0.46$) than the grades given by female professors ($M = 6.52$; $SD = 1.04$). Meanwhile, in the male majors, there were no significant differences in the means of the participants based on the professor's gender, $t(146) = 0.89$, $p = .375$, $d = 0.07$.

The ANOVA also showed a significant two-way interaction between the variables participant's gender and professor's gender, $F(1,322) = 11.347$, $p < .01$, $\eta^2 = .035$. There was a significant difference in the evaluation means when the subject was taught by a female professor. The planned comparisons indicate that male students perceived the grades assigned by the female professors as less fair ($M = 6.20$; $SD = 1.28$) than the female students ($M = 6.65$; $SD = 0.78$), $t(51.98) = -2.10$, $p = .04$, $d = 0.29$. When the professor was male, there was no significant difference in the student's evaluation of the grade received, $t(175) = 0.645$, $p = .52$, $d = 0.05$. However, in this condition, the means of both the female and male students were very close to the maximum point of the scale used (7), reaching a plateau due to a ceiling effect that the grade attributed by the male professor was fair, but not when a female professor gave the grade.

Although the effect of the three-way interaction between the type of major, the participant's gender, and the professor's gender did not obtain statistical significance, $F(1, 322) = 1.626$, $p = .203$, $\eta^2 = .005$, the planned comparisons showed some significant effects between these variables when combined with each other. This analysis strategy is named as simple effects using one degree-of-freedom planned contrast (Judd et al., 1995).

Figure 1 illustrates that only the differences between the means of male students from majors considered female were significant, indicating that they perceived the grade given by a female professor of these majors as less fair ($M = 5.80$;

Figure 1

Perception of Fairness of the Best Grade depending on the Type of Major, Participant's Gender, and Professor's Gender

SD = 1.87) than when the grade was given by a male professor (*M* = 6.92; *SD* = 0.28), *t*(14.68) = 2.265, *p* = .039, *d*= 1.01. The means of the female students did not differ statistically, indicating that, for them, the professor's gender in the typically female majors did not influence the perceived fairness of the grade received, *t* (145) = 0.882, *p* = .379, *d* = 0.27

In the male majors, the planned comparisons did not show statistically significant differences in the condition of female students, *t* (49) = -0.696, *p* = .49, *d* = 0.10. However, in male majors, male students tended to perceive the grades given by male professors (*M* = 6.69, *SD* = 0.46) as fairer than those of female professors, (*M*= 6.42, *SD*= 0.70), *t* (95) = 2.16, *p* = .033, *d*= 0.45.

Discussion

Taken together, the results presented so far showed that grades attributed by male professors were perceived as fairer than those attributed by female professors. The tendency to evaluate male professors' performance as fairer than female professors was observed even in majors perceived as typically female, with particular note that male students rated the female professor worse than the male professor. Therefore, discrimination against women in the workplace does not necessarily stem from counter-stereotypical behavior (Bareket et al., 2021; Costa et al., 2023; González et al., 2019; Heilman et al., 2024; Nyul et al., 2025; Viana et al., 2020), given that discrimination occurred in both typically male and typically female majors. These results can be discussed based on the male supremacy ideology (Eagly & Wood, 2012; Heilman, 2012). Because men are perceived as more suited to professions involving leadership, decision-making power, and agentic tendencies (Eagly et al., 2000; Laksono & Patriot, 2021), the more favorable behavior toward men than women persists even in academic disciplines associated with the female gender.

Although no differences were found in the evaluation of male and female professors based on typically male or typically female majors, it is important to note that, in the present Study, participants were asked to evaluate professors from majors in which they had received the highest grade. Considering that job performance tends to influence professional evaluations — and that, in the case of academic careers, the grade assigned by the professor often affects how students assess the professor's performance (Palazzo & Gomes, 2009; Valentín-Martínez & Mayor-Ruiz, 2023) — it is worth considering whether the professor's performance could exacerbate gender-based discrimination in the workplace. More precisely, could the professor's performance result in differential evaluations of male and female professors across majors perceived as typically masculine versus those perceived as typically feminine? To answer this question, Study 2 was designed to determine the extent to which the professor's performance might intensify discriminatory practices. For this purpose, the professor's performance was defined at two levels (positive and negative), operationalized as feedback on an academic task completed by the student.

Study 2 – Evaluation of professors according to the grade received

This experimental study aimed to examine the relationship among the following variables: type of academic major, professor's gender, participant's gender, and the professor's performance—operationalized as the feedback (positive or negative) provided on an academic task completed by the student—regarding participants' perceptions of the fairness of the feedback received in typically male and typically female majors. The study was guided by the following hypotheses:

a) Given that professors tend to be evaluated more negatively by students when they assign low grades (Valentín-Martínez & Mayor-Ruiz, 2023), it was hypothesized that participants would perceive feedback as fairer when it is positive rather than negative (Hypothesis H1);
b) Considering gender disparities in the labor market (World Economic Forum, 2024), as well as the results of Study 1—which indicated that female professors are evaluated more negatively than male professors, regardless of whether the major is typically masculine or typically feminine—it was hypothesized that feedback provided by female professors would be perceived as less fair than feedback provided by male professors (Hypothesis H2);
c) Given the persistence of gender stereotypes in society—which associate men with agentic, assertive, and

independent traits, and women with caregiving and domestic responsibilities (Bareket et al., 2021; Costa et al., 2023; Eagly et al., 2000; González et al., 2019; Heilman et al., 2024; Nyul et al., 2025; Viana et al., 2020)—it was hypothesized that there would be an interaction among the type of major (typically male- vs. typically female), the professor's gender, and the type of feedback provided. Specifically, in typically male majors, female professors were expected to receive lower evaluations regardless of whether they gave positive or negative feedback (Hypothesis H3).

Method

Participants

This experimental study included 376 university students, of which 51.3% were students in typically male majors (engineering, physics, and computer Science), and 48.7% were students in majors considered female (pedagogy and psycho-pedagogy). Ages ranged from 17 to 62 (M = 23.4; SD = 6.51). Sample size was previously calculated using G*Power software, estimating a sample of 237 participants as adequate for an experimental design with 80% power to detect a moderate effect size between the variables (f = .25). However, just as in the previous study, we sought to collect as many participants as possible, within the concession granted by the universities

Instruments

An online questionnaire was used that presented to the participants situations specific to their major (typically male major or typically female major). Participants in the male majors answered questions about logical-mathematical reasoning, and those in the female majors solved questions related to teaching children with learning disabilities. The experimental manipulation was structured as follows: After completing a course-compatible task, participants received a grade from a supposed professor. The professor's gender was randomly assigned (via the SurveyMonkey program) to one of three conditions: Female Professor, Male Professor, or Control Condition (no gender mentioned), who was a specialist in the questions proposed in the situations presented. The feedback provided also included a randomized manipulation (implemented via the SurveyMonkey program), which varied across two levels: Positive Feedback or Negative Feedback. Each participant received only one type of feedback. As a dependent variable, participants were required to report how fair they considered the received feedback, using a Likert scale ranging from 1 (very unfair) to 7 (very fair). This variable was called the perceived fairness of the evaluation received.

Procedures

Data Collection. Participants independently completed the online questionnaire in a room with internet access. Data collection was carried out using the SurveyMonkey program, which ensured the randomization of the assessment carried out by the supposed major.

Data Analysis. Given that a four-way interaction analysis would present issues of parsimony and interpretability, due to the high number of interactions involved, as well as increase the likelihood of Type I error (Maxwell & Delaney, 2004), it was decided to conduct three separate Analyses of Variance (ANOVA) using a 2 × 3 × 2 design, performed with SPSS software, version 22:

a) an ANOVA taking as Independent Variables: Participant's Gender (Male x Female) X Professor's Gender (Male x Female x Control Condition) X Type of major (Male x Female), and as Dependent Variable, how fair they considered the received feedback
b) an ANOVA taking as Independent Variables, the type of feedback received (Positive x Negative) X Professor's Gender (Male x Female x Control Condition) X Participant's Gender (Male x Female), and as Dependent Variable, how fair they considered the received feedback;
c) an ANOVA taking as Independent Variables, the type of feedback received (Positive x Negative) X Professor's Gender (Male x Female x Control Condition) X Type of major (Male x Female), and as Dependent Variable, how fair they considered the received feedback.

Ethical Considerations

This work was approved by the Research Ethics Committee of Universidade Federal da Paraíba (Brazil) under the number 036/15. CAAE: 410228.0.0000.5188

Results

The ANOVA results showed a significant main effect only for the feedback type, F (1,375) = 34.427, p < .001, η^2= .089, indicating that the participants generally perceived the positive evaluations (M = 5.37, SD = 2.1) as being fairer than the negative evaluations (M = 3.8, SD =2.9). This corroborates the first hypothesis of this study (Hypothesis H1). The professor's gender also had a significant effect, F (2, 375) = 3.135, p = .045, η^2 = .018. Female professors (M = 4.35, SD = 1.93) were rated more negatively than male professors (M = 4.92, SD = 1.76), t (248) = 2.417, p = .016, d = 0.31, thus confirming Hypothesis H2. The Male professors also received higher ratings compared to the control condition (M = 4.57, SD = 1.82), t (250) = 1.512, p = .132, d = 0.19. No significant differences were observed between the female professor and control conditions, t (346) = −0.944, p = .346, d = −0.19. The main effects of the variables major type, F (1,375) = 1.634, p = .202, η^2= .005 and participant's gender, F (1,375) = .887, p = .347, η^2=.003 were not significative. Nevertheless, the planned comparisons found that some combinations between these variables had statistically significant differences; therefore, they should be analyzed as simple effects using a degree-of-freedom planned contrast (Judd et al., 1995).

When analyzing the planned comparisons, in an interaction between the type of major and the professor's gender, although not significant, F (2, 375) = 0.854, p = .426, η^2 = .005, differences were observed between the recorded means only in the major considered "male." This result showed differences between the perceived fairness of the grades given by male (M = 4.93, SD = 1.58) and female professors (M = 4.31, SD = 1.90), t (128) = 2.038, p = .04, d = 0.36. However, no differences were found in the recorded means in the control situation (without the professor's gender), in comparison to male professor, t (120.54) = 1.69, p = .09, d= 0.31 and female professor, t (125) = -0.295, p = .768, d= 0.05. Thus, in the typically male majors, the participants perceived the male professor as fairer than the female professor, corroborating the first part of our third hypothesis. In typically female majors, no differences were found when comparing male and female professors, t (118) = 1.39, p = .166, d = 0.26; between male professors and the control condition, t (121) = .465, p = .64, d = 0.08; or between female professors and the control condition, t (121) = -1.03, p = .30, d = 0.19.

The interaction between the professor's gender and type of feedback (Figure 3) was not significant, F (2, 375) =0.468, p = .627, η^2 = .003. As can be seen in Figure 2, these results show that positive evaluations are seen as fairer than negative ones, regardless of the gender of the professor who gives the feedback, since there were significant differences even in the means of the control situation. Male Professor: Positive feedback (M =5.6 SD =1.25), Negative feedback (M= 4.19, SD =1.93), t(124) = 4.86, p < .001, d = 0.87, Female Professor: Positive feedback (M = 5.19, SD = 1.52), Negative feedback (M= 3.51, SD = 1.95), t(115.1) = 5.32, p < .001, d= 0.96; Control Condition: Positive feedback (M =5.18, SD = 1.52), Negative feedback (M = 3.91, SD = 1.89), t(124) = 4.137, p< .001, d = 0.74. However, for the situation in which a female professor gives the negative feedback, the means are very close to the midpoint of the scale, which may indicate discrimination against the female professor, who is supposedly more rigid and provides a negative evaluation.

The three-way interaction between the variables participant's gender, type of major, and professor's gender was not significant, F (2,375) = .006, p = .994, η^2 <.001. Significant differences, nevertheless, were found between male and female students in typically female majors. In these majors, female students rated the female professor as less fair (M = 4.10, SD = 2.00) than male students did (M = 5.58, SD = 1.37), t (58) = 2.40, p = .020, d = 0.88. This finding shows that women in the typically female context perceive feedback by female professors as more unfair. In typically female majors, no differences were found between male and female students' evaluations of the male professor, t (58) = 0.54, p = .584, d = 0.14. Participant gender also did not produce significant differences in evaluations of the control condition, t (61) = 0.88, p = .384, d = 0.22. In typically male majors, there were no gender-based differences in evaluations of: the male professor, t (64) = -1.24, p = .220, d = 0.31; the female professor, t (62) = 0.66, p = .513, d = 0.17; or the control condition, t (61) = -0.45, p = .652, d = 0.11

Figure 3 shows the three-way interaction between the participant's gender, the professor's gender, and the type of feedback, F (2, 375) = 2.170, p = .116, η^2 =.01. The results indicate that the differences lie in how male and female participants evaluate the feedback received from the female professor. Although the participant's gender did not influence the evaluation of the female professor when she provided positive feedback, t (60) = -0.69, p = .48, d = 0.18, differences were found in the condition where the professor gave negative feedback. In this condition, female participants rated the female

Figure 2

Perceived Fairness of Feedback as a Function of the Professor's Gender and the Type of Feedback

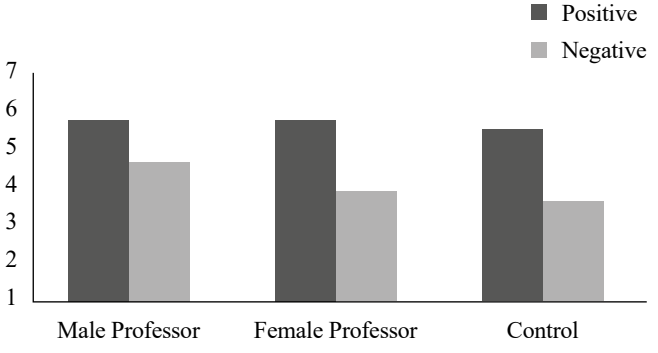
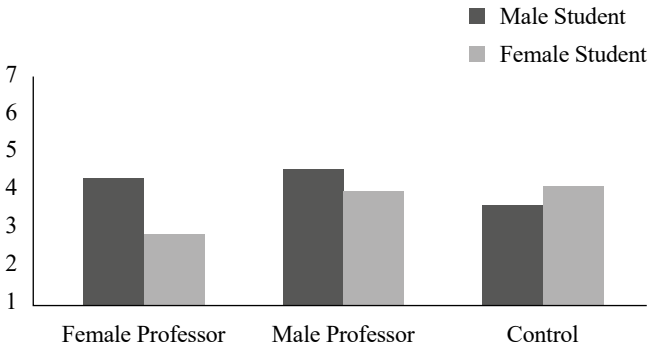


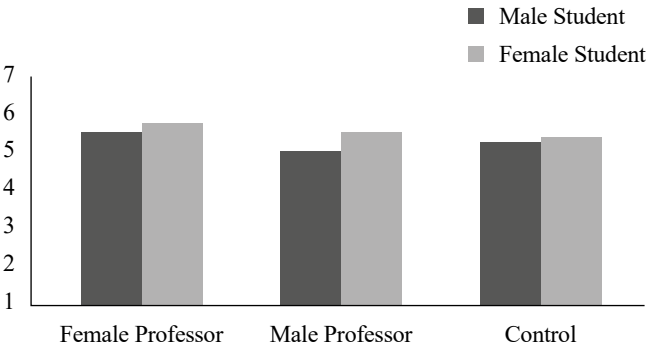
Figure 3

Perceived fairness of the Feedback received according to the Professor's Gender and Type of Feedback

Negative Feedback



Positive Feedback



professor (M = 2.87, SD = 1.83) more negatively than male participants (M = 4.16, SD = 1.88), t (60) = 2.731, p = .008, d = 0.70. In other words, regardless of the major, the female students discriminate more against the feedback provided by their female professors. This result, although not foreseen in the hypotheses, seems to be that discrimination does not only come from men against women. The remaining conditions did not produce significant differences. The participants' gender did not influence the evaluation of the male professor, either when he provided positive feedback, t (63) = 1.56, p = .12, d = 0.39, or when he provided negative feedback, t (59) = 0.728, p = .46, d = 0.19. No differences were observed in the control condition, whether it involved positive feedback, t (64) = -0.36, p = .71, d = 0.09, or negative feedback, t (58) = 0.70, p = .48, d = 0.18.

The three-way interaction between the type of major, professor's gender, and type of feedback was significant, F (2, 375) = 3.239, p = .04, η^2 = .01 (Figure 4). Thus, in male majors, differences were found in the means both in the situation in which feedback was given by a male professor: Positive Feedback (M = 5.42, SD = 1.33), Negative Feedback (M = 4.38, SD= 1.68), t (64) = 2.79, p = .007, d = 0.70, and the control condition, Positive Feedback (M = 5.08, SD = 1.63), Negative Feedback (M=3.62 SD = 1.93), t (61) = 3.25, p = .002, d = 0.82, however, no differences were observed in the evaluation of the female professor based on whether she provided positive or negative feedback, t (62) = 1.72, p = .08, d = 0.69. As shown in Figure 4, the teacher's ratings clustered near the midpoint of the scale regardless of the type of feedback she provided.

Taken together, the results once again revealed that participants in the typically male-major rated the professor's feedback as fairer when it was positive compared to negative. However, this pattern did not hold for the female professor, who received lower evaluations regardless of whether she provided positive or negative feedback.

In female majors, the means were also different in the conditions in which a male professor gave feedback, Positive Feedback (M = 5.0, SD = 1.63), Negative Feedback (M = 3.62, SD= 1.9), t (61) = 3.259, p = .002, d = 0.78, and in the situation in which the female professor gave feedback, Positive Feedback (M = 5.7, SD = 1.1), Negative Feedback (M = 3.1, SD= 1.80) t(58) = 6.7, p <.001, d= 1.74, and the control condition, Positive Feedback (M = 5.28, SD = 1.41), Negative Feedback (M = 4.19, SD= 1.85), t(61) = 2.62, p = .011, d = 0.66. Thus, this result indicates that, contrary to what occurs in typically male majors, positive feedback is perceived as fairer than negative feedback, both when it is given by a male professor and when it is given by a female professor.

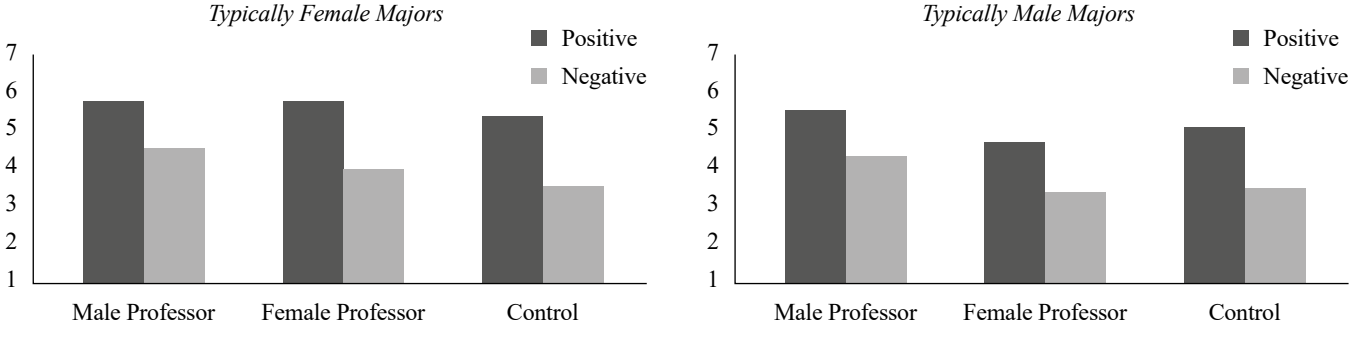
These findings suggest an interaction between the type of major (typically male vs. typically female), the professor's gender, and the feedback provided, and they partially confirm Hypothesis H3. Specifically, in typically male-dominated fields, although the evaluation of the female professor is not necessarily negative, since it approximates the midpoint of the scale, the fact that she provides positive feedback to the participant is not sufficient for her evaluation to be perceived as significantly fairer than in the negative feedback condition. The same does not hold for the male professor, who is evaluated more favorably when his feedback is positive.

Discussion

The results of Study 2 indicate that while the type of feedback (positive or negative) was the most influential factor in students' perceptions of fairness, the professor's gender also played a role, particularly depending on the professional context (typically male or typically female). This was especially evident in the lower average ratings of positive feedback when it was provided by female professors in typically male majors. These results make two important contributions. First, they confirm that discrimination against women — here observed through more favorable evaluations of male professors — is an aspect also verified in Study 1. Second, the results also indicate that discrimination against the female professor is exacerbated by the interaction between the typically male-dominated work context and her workplace performance, here operationalized through the feedback (positive vs. negative) provided to the student—a finding that aligns with previous research pointing to less favorable treatment of women based on their job performance, and with the perception that careers in Science, Technology, Engineering, and Mathematics (STEM) are more appropriate for men (Bareket et al., 2021; Costa et al., 2023; González et al., 2019; Heilman et al., 2024; Laksono & Patriot, 2021; Nyul et al., 2025; Viana et al., 2020).

Figure 4

Perceived fairness according to the Type of Major, Professor's Gender, and Type of Feedback



One particularly striking result is that female students enrolled in typically male majors were those who most strongly discriminated against female professors when negative feedback was given. The female professor was even more punished by the female students, which also occurs in contexts perceived as female. These findings diverge from those obtained in Study 1, where male students provided more negative evaluations of the female professor. The inconsistency between findings calls for closer scrutiny in future studies.

One possible reason for this variation in outcomes may lie in the role played by belief systems. It is well-established that discriminatory behavior varies according to individuals' adherence to belief systems that legitimize social inequality. In the context of gender relations, Ambivalent Sexism and Social Dominance Orientation have been consistently identified as contributing to discrimination against women (Bareket & Fiske, 2023; Sidanius et al., 2024), shedding light on psychosocial processes that make discrimination appear legitimate — even among minority groups.

Since the present study did not measure these factors, future research should investigate how the interplay between gender, performance, work context, and belief systems contributes to the manifestation of discrimination.

General Discussion

The results of both studies indicate that gender is associated with how students perceive the fairness of grades assigned by professors. Moreover, this association appears to be stronger depending on the type of professional context involved. These results allow reflection on gender inequality in university teaching and indicate, among other things, that, despite the advances achieved by women in the labor market, they are still perceived as different from men, especially when they decide to enter contexts considered "male." In addition, these results also prompt reflection on the current evaluation processes of professional performance. As demonstrated, the professor's gender and other variables relate to how students perceive their professors. For example, when a higher education institution systematically evaluates its instructors through student feedback without considering if the professors are male or female working in a context perceived as "male" or "female," this may bias performance evaluations and lead to serious and concerning consequences for instructors' careers and psychological well-being.

Thus, the studies presented here could greatly contribute to how professor evaluations have been carried out by showing the need to consider macro-social aspects such as gender and professional context for taking positions in this area. These studies are also important for discussions about gender inequality in university teaching, because our results reveal that the grades attributed by female professors were often evaluated as being more unfair than those given by male professors, even when both had the same attributes and competencies. In other words, the results seem to demonstrate that even on equal terms, men are still given the benefit of the doubt and probably perceived as more competent than women.

However, future studies could conduct an experiment that can directly measure the impact of gender on professional evaluation, since, in this article, the operationalization of the dependent variable was done indirectly through the perceived fairness of the grades attributed by the professors. Although our study evidenced differences in the perceived fairness in experimentally controlled situations, we have no way of knowing the magnitude of the association of gender on the professional evaluation of university professors, as a study needs to consider several other variables in which gender is just one of them.

Therefore, future studies could include the main variables already used to carry out professional evaluation and include gender to compare which variables best predict professional performance, including other contexts of work activity in addition to university teaching. This study aimed to answer three questions: "Does the professor's gender in the academic environment impact the evaluations given to them?", "Is the type of professional context (e.g., type of major) in which these professors teach also related to these assessments?", and "Does the evaluation that a professor gives to the student also impact their evaluation of the professor?" To answer them, two main studies were carried out, all with university students who evaluated the grades attributed and evaluations made by their respective professors.

Regarding the first question, the results of Study 1 indicate that, together, the grades by male professors were seen as fairer by both female and male students. In addition, it is important to keep in mind that the means were very close to the maximum point of the scale, pointing to a consensus in relation to the professor's performance. Regarding the second question, together, the results of Study 1 show that, in female majors, when a male teacher taught the subject, the grade he attributed is seen as being fairer. Meanwhile, in the male majors, there were no significant differences in the means of the participants. In addition, in female majors, male students perceive the grades assigned by female professors as less fair than those given by male professors. However, in these majors, for the female students, the professor's gender did not relate to the perception of the grade received. The male majors also showed a tendency toward consensus: for both male and female students, the grades attributed by male professors were perceived as fairer than those given by female professors. The results of Study 2 show that, in male majors, the participants generally perceived the male professor as fairer than the female professor.

Finally, Study 2 was designed to specifically answer the third question, and together, the results indicate that positive evaluations were perceived to be fairer than negative evaluations. However, when the female professor provides negative feedback, she is perceived as even more unfair than the male professor, even without considering the type of major.

Taken together, the results presented here demonstrate that the professor's gender, the type of major taught, the evaluation given, and, to a lesser extent, the student's gender relate to the perception that the student body makes of its teaching faculty.

These results indicate a general consensus about the perceived fairness of grades given by male professors. The same is not true for female professors. For them, the perceived fairness of the evaluation was more easily negatively influenced by the factors investigated here.

These results raise some questions both in relation to performance evaluations of instructors and the sexism that still exists in university teaching. Thus, from the information presented here, on the one hand, intervention strategies emphasize that both men and women can practice any type of profession. On the other hand, performance evaluations must present strategies that mitigate the intervening variables that can influence them. Another limitation concerns the measure of the dependent variable. In the studies conducted here, the dependent variable was measured by only one item, which limits the possibility of obtaining greater reliability in the replication of the results. Thus, further research is needed to develop more robust measures for the assessment of discriminatory conduct against women in the workplace.

Finally, in addition to the need for future studies to consider the important role played by belief systems (Bareket & Fiske, 2023; Sidanius et al., 2024), we can identify other gaps that can be explored in future work. The first concerns the adoption of more direct strategies for evaluating teaching performance, such as mastery of the content and clarity in their presentation. Thus, they could analyze in more detail which dimensions are more specific according to the professor's gender. A second gap left by the studies presented here is related to the fact that we use gender was treated in a binary manner. This view is reductionist, and gender, in its more diverse forms of expression, would be more appropriate (Cartwright & Nancarrow, 2022). However, we opted for simplification since it is the first work that analyzes the issue of discrimination against women in the workplace, taking into account the aspects investigated here.

Data Availability

The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

References

Bareket, O., & Fiske, S. T. (2023). A systematic review of the ambivalent sexism literature: Hostile sexism protects men's power; benevolent sexism guards traditional gender roles. *Psychological Bulletin, 149*(11-12), 637-698. https://doi.org/10.1037/bul0000400

Bareket, O., Shnabel, N., Kende, A., Knab, N., & Bar-Anan, Y. (2021). Need some help, honey? Dependency-oriented helping relations between women and men in the domestic sphere. *Journal of Personality and Social Psychology, 120*(5), 1175–1203. https://doi.org/10.1037/pspi0000292

Basow, S. A. (2000). Best and worst professors: Gender patterns in students' choices. *Sex Roles*, *43*(5-6), 407-417. https://doi.org/10.1023/A:1026655528055

Basow, S. A., Phelan, J. E., & Capotosto, L. (2006). Gender patterns in college students' choices of their best and worst professor. *Psychology of Women Quarterly*, *30*(1), 25–35. https://doi.org/10.1111/j.1471-6402.2006.00259.x

Cartwright, T., & Nancarrow, C. (2022). A question of gender: Gender classification in international research. *International Journal of Market Research, 64*(5), 575-593. https://doi.org/10.1177/14707853221108663

Connor, R. A., & Fiske, S. T. (2019). Not minding the gap: How hostile sexism encourages choice explanations for the gender income gap. *Psychology of Women Quarterly*, *43*(1), 22–36. https://doi.org/10.1177/0361684318815468

Cortina, C., Rodríguez, J., & González, M. J. (2021). Mind the job: The role of occupational characteristics in explaining gender discrimination. *Social Indicators Research, 156*, 91–110. https://doi.org/10.1007/s11205-021-02646-2

Costa, L. L. A., Viana, H. A., & Albuquerque, I. M. (2023). Gênero [Gender]. In A. R. R. Torres, M. E. O. Lima, E. M. Techio, & L. Camino (Eds.), *Psicologia social: Temas e teorias* [Social psychology: Themes and theories] (pp. 445-464). Edgard Blücher.

Eagly, A. H., & Wood, W. (2012). Social role theory. In P. A. M. Van Lange, A. W. Kruglanski, & E. T. Higgins (Eds.), *Handbook of theories in social psychology* (Vol. 2, pp. 458–476). Sage.

Eagly, A. H., Wood, W., & Diekman, A. B. (2000). Social role theory of sex differences and similarities: A current appraisal. In T. Eckes & H. M. Trautner (Eds.), *The developmental social psychology of gender* (pp. 123–174). Lawrence Erlbaum.

Expósito, F., Moya, M. C., & Glick, P. (1998). Sexismo ambivalente: Medición y correlatos [Ambivalent sexism: Measurement and correlates]. *Revista de Psicologia Social*, *13*(2), 159-169. https://doi.org/10.1174/021347498760350641

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (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*(6), 878–902. https://doi.org/10.1037/0022-3514.82.6.878

Glick, P., & Fiske, S. T. (1996). The ambivalent sexism inventory: Differentiating hostile and benevolent sexism. *Journal of Personality and Social Psychology*, *70*(3), 491-512. https://doi.org/10.1037/0022-3514.70.3.491

González, M. J., Cortina, C., & Rodríguez, J. (2019). The role of gender stereotypes in hiring: A field experiment. *European Sociological Review*, *35*(2), 187–204. https://doi.org/10.1093/esr/jcy055

Hanek, K. J., & Garcia, S. M. (2022). Barriers for women in the workplace: A social psychological perspective. *Social and Personality Psychology Compass, 16*(10), e12706. https://doi.org/10.1111/spc3.12706

Heilman, M. E. (2012). Gender stereotypes and workplace bias. *Research in Organizational Behavior, 32*, 113–135. https://doi.org/10.1016/j.riob.2012.11.003

Heilman, M. E., Caleo, S., & Manzi, F. (2024). Women at work: Pathways from gender stereotypes to gender bias and discrimination. *Annual Review of Organizational Psychology and Organizational Behavior, 11*, 165–192. https://doi.org/10.1146/annurev-orgpsych-110721-034105

Hideg, I., & Krstic, A. (2021). The quest for workplace gender equality in the 21st century: Where do we stand and how can we continue to make strides? *Canadian Journal of Behavioural Science, 53*(2), 106–113. https://doi.org/10.1037/cbs0000222

Instituto Brasileiro de Geografia e Estatística. (2022). *Pesquisa nacional por amostra de domicílios contínua: 4º trimestre de 2013 e de 2022* [Continuous national household sample survey: 4th quarter of 2013 and 2022]. https://sidra.ibge.gov.br/tabela/5436#resultado

Judd, C. M., McClelland, G. H., & Culhane, S. E. (1995). Data analysis: continuing issues in the everyday analysis of psychological data. *Annual Reviews of Psychology*, *46*, 433-465. https://doi.org/10.1146/annurev.ps.46.020195.002245

Laksono, P. J., & Patriot, E. A. (2024). Exploring gender dynamics in STEM education: A literature review. *International Conference on Research Issues and Community Service, 1*, 140–146. https://doi.org/10.31332/i-cores.v1i1.11070

Lei No. 14.611. (2023, 3 de julho). Dispõe sobre a igualdade salarial entre mulheres e homens [Provides for equal pay between women and men]. https://www.planalto.gov.br/ccivil_03/_ato2023-2026/2023/lei/l14611.htm

Liazibra, L. F. (2024, March 8). *Lei de igualdade salarial entre homens e mulheres foi vitória histórica* [Equal pay law for men and women was a historic victory] [Speech audio recording]. Rádio Senado. https://www12.senado.leg.br/radio/1/noticia/2024/03/08/lei-de-igualdade-salarial-entre-homens-e-mulheres-foi-vitoria-historica

MacNell, L., Driscoll, A., & Hunt, A. N. (2015). What's in a name: Exposing gender bias in student ratings of teaching. *Innovative Higher Education, 40*, 291-303. https://doi.org/10.1007/s10755-014-9313-4

Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: A model comparison perspective* (2nd ed.). Lawrence Erlbaum.

Nyul, B., Ksenofontov, I., Fleischmann, A., & Kahalon, R. (2025). Brilliance as gender deviance: Gender-role incongruity as another barrier to women's success in academic fields. *Journal of Experimental Social Psychology, 116*, 104680. https://doi.org/10.1016/j.jesp.2024.104680

Palazzo, J., & Gomes, C. A. (2009). Plano de carreira e avaliação dos professores: Encontros e desencontros [Career plan and teacher evaluation: Agreements and disagreements]. *Ensaio: Avaliação e Políticas Públicas em Educação*, *17*(63), 205-234.

Picanço, F., Araújo, C. M. O., & Covre-Sussai, M. (2021). Papéis de gênero e divisão das tarefas domésticas segundo gênero e cor no Brasil: Outros olhares sobre as desigualdades [Gender roles and division of household tasks according to gender and race in Brazil: Other perspectives on inequalities]. *Revista Brasileira de Estudos de População, 38*, 1-31. https://doi.org/10.20947/S0102-3098a0177

Sidanius, J., Hudson, S.-K. T. J., Davis, G. K., & Bergh, R. (2018). The theory of gendered prejudice: A social dominance and intersectionalist perspective. In A. Mintz & L. G. Terris (Eds.), *The Oxford handbook of behavioral political science* (pp. 253-284). Oxford Academic.

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

Tabassum, N., & Nayak, B. S. (2021). Gender stereotypes and their impact on women's career progressions from a managerial perspective. *IIM Kozhikode Society & Management Review, 10*(2), 192-208. https://doi.org/10.1177/2277975220975513

Valentín-Martínez, B., & Mayor-Ruiz, C. (2023). La evaluación del profesorado universitario: Una mirada del estudiantado. *Educação e Pesquisa, 49*, e241907. https://doi.org/10.1590/S1678-4634202349241907esp

Viana, H. A., Rosas-Torres, A. R., Estramiana, J. L. A. (2020). Egalitarian men: Stereotypes and discrimination. *Acta Colombiana de Psicología*, *23*(2), 111-147. http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0123-91552020000200111

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

Hyalle Abreu Viana is a Professor at Centro Universitário Maurício de Nassau, João Pessoa - PB, Brazil.

Khalil da Costa Silva is a Professor at the Instituto de Psicologia e Serviço Social of the Universidade Federal da Bahia, Brazil.

Elza Maria Techio is a Professor at the Instituto de Psicologia e Serviço Social of the Universidade Federal da Bahia, Salvador-BA, Brazil.

José Luis Álvaro is a Professor in the Departamento de Psicología Social at the Universidad Complutense de Madrid, Somossaguas, Madrid, Spain.

Ana Raquel Rosas Torres is a Professor in the Departamento de Psicologia at the Universidade Federal da Paraíba, João Pessoa – PB, Brazil.

Authors' Contribution:
All authors made substantial contributions to the conception and design of this study, to data analysis and interpretation, and to the manuscript revision and approval of the final version. All the authors assume public responsibility for the content of the manuscript.

Associate editor:
Cicero Roberto Pereira

Received: Feb. 12, 2025
1st Revision: May 15, 2025
2nd Revision: Jul. 19, 2025
Approved: Jul. 24, 2025

How to cite this article:
Viana, H.A., da Costa Silva, K., Techio, E.M., Álvaro, J.L., & Torres, A. R. R. (2025). Sexism in University Teaching: Differences in the Evaluations of Male and Female Professors. *Paidéia (Ribeirão Preto), 35,* e3531. https://doi.org/10.1590/1982-4327e3531