RUCH PRAWNICZY, EKONOMICZNY I SOCJOLOGICZNY
Rok LXXXVIII – zeszyt 2 – 2026
https://doi.org/10.14746/rpeis.2026.88.2.16

PATRYCJA JANOWSKA-WIDOMSKA,[a] MICHAŁ KRAWCZYK,[b]
DOMINIKA MAISON[c]

CAN A PROFESSOR BE A WOMAN? GENDERED PERCEPTIONS IN A RIDDLE EXPERIMENT

CZY PROFESOR MOŻE BYĆ KOBIETĄ? EKSPERYMENT OPARTY NA ZAGADCE

This preregistered study examined how unconscious gender schemas, demographic characteristics, and life experiences influence the recognition of women in academic authority roles. A sample of US-based students (n = 1,686) was presented with a riddle task modelled after the classic surgeon scenario, adapted to a university setting. Solving the riddle required identifying a professor as the child's mother, thereby challenging traditional gender schemas. Only 17 per cent of participants arrived at this interpretation, despite clear narrative cues. The professor's disciplinary affiliation – whether a traditionally male-dominated field (physics), a female-dominated field (art history), or a gender-balanced field (geography) – had little effect on solution rates. We conducted logistic regression analyses to assess whether exposure to female academics, political orientation, feminist identity, or other demographic factors predicted successful identification of the professor as a woman. Interestingly, male participants were more likely than female participants to solve the riddle in this way. This unexpected finding may reflect differences in schema activation or response strategies, although further investigation is needed to clarify the underlying mechanisms. Notably, willingness to donate to a foundation supporting women in academia – a measure of prosocial behaviour – did not significantly predict the proposed solution to the riddle. A supplementary study supported the face validity of the task: participants judged the 'mother' solution to be both plausible and typically overlooked due to prevailing stereotypes. These findings demonstrate that gender schemas remain deeply embedded and can shape cognitive interpretations even when contextual cues contradict them. The results underscore the persistence of implicit bias in academic contexts.

Keywords: gender schemas; implicit association; role schemas; gender roles; stereotyping

W tym preregistrowanym badaniu sprawdzono, w jaki sposób nieuświadomione schematy płciowe, cechy demograficzne oraz życiowe doświadczenia wpływają na rozpoznawanie kobiet jako autorytetów w środowisku akademickim. Uczestnikom – studentom z USA (n = 1686) – przedstawiono zagadkę inspirowaną klasyczną historią o chirurgu, dostosowaną do kontekstu uniwersyteckiego.

[a] University of Warsaw, Poland / Uniwersytet Warszawski, Polska
p.janowska2@uw.edu.pl, https://orcid.org/0000-0002-0853-2170
[b] University of Warsaw, Poland / Uniwersytet Warszawski, Polska
mkrawczyk@wne.uw.edu.pl, https://orcid.org/0000-0001-6795-1175
[c] University of Warsaw, Poland / Uniwersytet Warszawski, Polska
dominika.maison@psych.uw.edu.pl, https://orcid.org/0000-0002-8946-7517

RPEiS 88(2), 2026: 267–288. *ISSN (Online) 2543-9170 ISSN (Print) 0035-9629*

Aby poprawnie ją rozwiązać, należało zidentyfikować profesora jako matkę dziecka – wbrew stereotypowym skojarzeniom dotyczącym płci. Tylko 17% badanych udzieliło tej odpowiedzi, mimo że wskazówki zawarte w historii wyraźnie na to wskazywały. Płeć przypisywana profesorowi nie była istotnie zależna od dyscypliny naukowej, niezależnie od tego, czy była to zmaskulinizowana dziedzina (fizyka), dziedzina zrównoważona pod względem płci (geografia) czy dziedzina sfeminizowana (historia sztuki). Analizy regresji logistycznej wykazały, że takie czynniki, jak kontakt z kobietami pracującymi naukowo, poglądy polityczne bądź identyfikacja feministyczna nie miały istotnego wpływu na poprawność odpowiedzi. Co zaskakujące, to mężczyźni częściej niż kobiety wskazywali na matkę jako rozwiązanie zagadki. Może to odzwierciedlać różnice w aktywacji schematów poznawczych lub stosowanych strategiach odpowiedzi, choć w celu stwierdzenia tego są wymagane są dalsze badania. Również deklarowana gotowość do wsparcia kobiet w nauce (poprzez darowiznę) nie przewidywała trafności rozwiązania. Dodatkowe badanie potwierdziło trafność zastosowanego narzędzia – uczestnicy uznali odpowiedź wskazującą matkę za wiarygodną, a jej pomijanie za efekt utrwalonych stereotypów. Wyniki wskazują, że schematy płciowe są głęboko zakorzenione i odporne na wpływ indywidualnych przekonań czy doświadczeń.

Słowa kluczowe: schematy płciowe; utajone skojarzenia; schematy ról; role płciowe; stereotypy

I. INTRODUCTION

Implicit or unconscious bias is frequently cited as a key contributor to women's underrepresentation in academia. Rooted in traditional gender roles, senior academic positions are often associated with men (see, among others, Di Tullio, 2019; Fan et al., 2019). Such bias may shape evaluations of female academics and hinder progress toward gender equity in academic institutions.

Measuring gender-biased thinking, however, is challenging, as individuals may be unaware of, or unwilling to report, such patterns. Researchers have therefore used indirect measures of bias, including riddles. One well-known example is the 'surgeon riddle', in which a father dies in a car crash and his son is taken to surgery, where the surgeon states, 'I can't operate on this boy – he is my son.' Respondents must determine how this is possible. Across multiple studies (Belle et al., 2021; Kollmayer et al., 2018; Morehouse et al., 2022; Stoeger et al., 2004), most participants fail to identify the surgeon as the boy's mother. As Reynolds et al. (2006) discuss, once a role is encoded as male, participants may struggle to revise that initial representation even when the narrative itself rules it out.

All prior studies using this paradigm have examined variants of the surgeon or specialist riddle. We extend this literature by shifting the setting to academia and introducing a 'professor' riddle. More importantly, rather than focusing on linguistic cues, our design tests whether gender bias varies across fields with different gender compositions. Specifically, we compare physics, geography, and art history to examine whether the prevalence of women in a discipline weakens the default association between professorship and men.

To do so, we developed a modified riddle in which a child arriving at a university receives a phone call from a professor who addresses the child as 'son' or 'daughter', implying that the professor is the other parent. This design

allows us to test whether difficulty in identifying the professor as female is stronger in male-dominated fields and weaker in female-dominated fields. We also examine whether the gender of the child affects responses. Evidence on this issue is mixed: Morehouse et al. (2022) found that gender-neutral wording ('child') reduced bias, whereas explicitly female wording ('daughter') was less effective.

Following Belle et al. (2021), we also examine whether women are more likely than men to identify the professor as the mother in the female professor condition, and whether greater exposure to female academics is associated with a higher likelihood of doing so. In addition, we extend the analysis beyond riddle performance by analysing open-ended responses, donation willingness, and implicit associations related to gender and academia.

Together, this study introduces a novel riddle paradigm for examining implicit gender stereotypes about professors across academic disciplines. By linking stereotype activation to cognitive responses and a behavioural measure, it contributes to the literature on gender bias in academia and extends prior riddle-based approaches to a new institutional context. To complement the riddle task, we also consider implicit associations and a supplementary face-validity assessment.

The remainder of the article is organized as follows. We first review the relevant literature and describe the methodology and hypotheses. We then present the results and discuss their implications for gender stereotypes in academia, including the methodological contribution of adapting the classic riddle paradigm to the academic context and enabling comparisons across disciplines. The final sections address the limitations of the study and conclude.

II. LITERATURE REVIEW

1. Gender stereotypes in academia

Academia is an environment in which gender-role schemas remain highly consequential. Although some disciplines have become more gender-balanced, and some are even female-dominated, academic authority and senior scholarly roles continue to be associated disproportionately with men (Kugler et al., 2021; Mastekaasa & Smeby, 2008). Experimental evidence shows that this bias shapes evaluations even when qualifications are held constant. Moss-Racusin et al. (2012), for example, found that science faculty rated female applicants as less competent and less hireable than otherwise identical male applicants, while also offering them lower starting salaries and less mentoring. Similarly, Steinpreis et al. (1999) reported that both male and female evaluators were more likely to recommend hiring a male candidate than an otherwise identical female candidate.

Such patterns are consistent with broader evidence that women are seen as less naturally suited to success in prestigious intellectual domains. Leslie et al. (2015) show that women are underrepresented in fields believed to

require innate brilliance, and Carli et al. (2016) argue that women are often perceived as lacking traits associated with scientific success. These findings suggest that academic competence and authority remain cognitively linked to men, especially in elite or high-status roles.

This tendency is also reflected in how scientists and faculty are perceived. A meta-analysis of Draw-A-Scientist studies shows that scientists are still predominantly depicted as men, despite some increase in female representations over time (Miller et al., 2018). Similar biases appear in university teaching contexts: male professors are often perceived as more competent, knowledgeable, and authoritative, and student evaluations tend to disadvantage women even when instructional quality is held constant (Boring, 2017; MacNell et al., 2015; Okoye et al., 2020; Rojek et al., 2019). Together, this literature suggests that the category of 'professor', much like that of 'scientist', remains strongly gendered.

2. Role congruity and gender schemas

These patterns can be understood through role congruity theory. According to this framework, prejudice arises when the attributes associated with a social group are perceived as inconsistent with the requirements of a valued role (Eagly & Diekman, 2005). In a related argument, Eagly and Karau (2002) propose that women face prejudice in leadership contexts because the female gender role is perceived as incongruent with leadership roles. As a result, women may be evaluated less favourably both as potential occupants of such roles and when enacting role-consistent behaviour.

This logic is particularly relevant in academia, where professorship combines authority, expertise, and leadership. If the role of professor is implicitly coded as male, then identifying a woman as the occupant of that role may be cognitively less accessible, especially under conditions requiring quick inference rather than explicit reflection. Gender schema theory similarly suggests that individuals rely on internalized gender categories to interpret social information (Bem, 1981). These schemas can persist even among individuals who consciously endorse egalitarian views, including women themselves (Rahmani Azad et al., 2023).

Prior work also suggests that internalized schemas may shape performance and self-perception. Belle et al. (2021), for example, show that even among college women with employed mothers and professional aspirations, stereotypical assumptions remain salient in riddle responses. This makes it plausible that female students may also hold the default association between professorship and men, even when they explicitly reject gender inequality.

3. Riddle paradigms and the present study

One way to capture such schemas indirectly is through riddle-based paradigms. The best-known example is the surgeon riddle, in which respondents often fail to identify the surgeon as the boy's mother despite the logic of the

story ruling out the father (Belle et al., 2021; Kollmayer et al., 2018; Morehouse et al., 2022; Stoeger et al., 2004). As Reynolds et al. (2006) argue, once a role is encoded as male, participants may struggle to revise that initial representation even when the narrative itself requires doing so. The riddle paradigm is therefore useful because it reveals automatic role-based assumptions in an indirect and open-ended way.

All prior studies using this paradigm have examined variants of the surgeon or specialist riddle. The present study extends this literature by shifting the setting to academia and introducing a professor riddle. This allows us to test whether professorship is similarly coded as male and, crucially, whether this bias varies across disciplines with different gender compositions. Unlike prior riddle studies, which mainly manipulated language cues, our design explicitly examines whether field feminization weakens the default association between professorship and men. To do so, we compare three fields: physics, geography, and art history, which differ in gender composition and cultural stereotypes.

Prior riddle research also informs our expectations about possible mechanisms and moderators. Kollmayer et al. (2018) show that linguistic cues can shape responses, which relates to our manipulation of whether the child is described as a son or daughter. Morehouse et al. (2022) find that gender-neutral wording reduces bias, suggesting that linguistic framing may affect the accessibility of non-stereotypical interpretations. Belle et al. (2021), in turn, examine individual-level predictors of solving the riddle and show that stereotypical assumptions can persist even among respondents with relatively egalitarian backgrounds. Building on this literature, we examine whether responses vary by participant gender and by exposure to female academics.

Finally, we extend the riddle paradigm beyond the immediate solution itself. In addition to open-ended responses, we consider whether bias revealed in the riddle is associated with donation willingness toward an organization combating gender stereotypes in academia. Because implicit attitudes may shape not only judgments but also behaviour, donation willingness provides an additional behavioural measure of stereotype-related responses (e.g., Fung et al., 2023).

III. METHODS

1. Outcome variable and response coding

Given the riddle's open-ended nature, the main question – 'How is this possible?' – does not have a single correct answer.[1] For the purposes of this

[1] Nonetheless, our riddle offers a narrower range of possible solutions compared to the surgeon riddle. We deliberately refined the surgeon riddle, partly to avoid certain alternative explanations that could arise there, and partly because the original scenario is extreme and potentially evokes strong emotional responses. For example, in the hospital scenario, a valid answer could be that the surgeon does not recognize the child because their face is unrecognizable due to the

study, we focus on responses that imply a two-parent heterosexual configuration in which one parent is in the car and the other is on the phone. We refer to such responses as 'normative' only as a shorthand for this focal response category. In this context, the term reflects the most socially conventional inference, given that in the large majority of families one parent is male and the other is female. Obviously, the use of this term does not imply any ethical judgment about any family arrangement, nor does it suggest that other interpretations are less valid. Responses involving same-sex couples (e.g. Belle et al., 2021) are therefore treated as valid alternatives, but they fall outside the focal category analysed here.

Responses were coded by a single coder (a co-author) using a simple rule-based procedure applied separately for each treatment condition. In the female professor condition, if a participant's response contained an explicit reference indicating that the professor was female (e.g. 'mother', 'mom', 'woman'), the response was coded as 'normative' = 1; all other responses were coded as 'normative' = 0. The same logic was applied in the male professor condition: if the response explicitly indicated that the professor was male (e.g. 'father', 'dad', 'man'), the response was coded as 'normative' = 1; otherwise it was coded as 'normative' = 0. Ambiguous responses (e.g. 'a parent', 'another parent') that did not indicate the professor's sex were thus coded as non-'normative', because they merely restated the structure of the puzzle without resolving it.

This coding scheme introduces an asymmetry: given the strong stereotype linking senior academic roles with men, identifying the caller as the mother in the female professor condition is more demanding than identifying the caller as the father in the male professor condition. We initially included both parent conditions in the first recruitment wave, while the preregistered second wave focused on the female professor condition. After the first wave, however, we observed that a notable share of responses in the male professor condition did not explicitly identify the father. To preserve a balanced comparison across parent conditions, we therefore extended the second wave to include both parent conditions as well. We report this post-preregistration adjustment here to maximize transparency.

2. Participants

The survey was distributed among American students. One group of participants consisted of students recruited via Prolific (399 participants). Another group consisted of students from Texas A&M University (TAMU) who were recruited through an email invitation to complete a survey. A total of 1,287 TAMU students participated. Before beginning the study, all participants reviewed and accepted an informed consent form (see Supplement 1[2]).

---

accident, or that the child brought in from the accident simply resembles the surgeon's child. Such possibilities do not apply in the case of our riddle.

[2] Supplements are available on the Open Science Framework, https://osf.io/gu36b/overview?view_only=196df620b412483bbda46991a64cd907

Participants recruited via Prolific received a small fixed reward for their time in addition to being entered into a drawing for a $50 prize. The demographics of the two groups were very similar, and we observed no significant differences in the distribution of dependent variables between Prolific and TAMU participants; we therefore report the results jointly. On average participants were 23 years old ($sd = 5.55$) 37.6% were male, 52% were non-Hispanic White, 19.6% were Asian, 18.98% were Hispanics and 4.2% were Black. The gender composition was similar across ethnic groups, and the proportions of participants were comparable across recruitment methods.

At the end of the questionnaire, we asked the respondents to take part in the Implicit Association Test (IAT), which measured the extent to which they associated men and women with the exact sciences and the humanities, respectively. For technical reasons, only TAMU students, and not those from Prolific, were invited to participate in the IAT after completing the survey. Since taking the IAT significantly increased the length of the survey, we allowed participants to decide after the first part whether they also wanted to take the IAT, in order to reduce the dropout rate. Due to the low response rate and the absence of any meaningful correlation with the primary outcomes, the results are not reported in the main article but are available in Supplement 2.

The text of the riddle used in our study was specifically created for this research. However, some participants could have encountered the aforementioned surgeon riddle or a similar puzzle. As foreseen in the preregistered data analysis plan, we thus asked participants if they had seen a similar riddle before. Those who had (453 out of 1,686 records) were allowed to complete the survey to the end and were paid accordingly or had a chance to receive remuneration, but their data are not included in further analysis.

3. Procedure

The study was preregistered on AsPredicted.org (#100489). The hypotheses and analyses reported below follow the preregistration document, with a few explicitly indicated exceptions.

In an online questionnaire, participants faced a riddle that we expected to be more difficult for participants relying more strongly on traditional gender schemas. The chief inspiration for our riddle was a version published by Mindspace in 2020 for International Women's Day as part of the Embrace Equity movement, although we introduced several substantial changes.

In a between-subjects design, participants were introduced to a riddle featuring a conversation between a professor and his or her child. They were asked to provide the first solution to the riddle that came to their mind. The riddle came in 12 versions following the same general structure (the PLACEHOLDERS are discussed below):

> A PARENT is about to bring his/her CHILD to an interview. SHE/HE is applying to be a DISCIPLINE major. Just as they arrive at the university's parking lot, the CHILD's phone rings. HE/SHE looks at HIS/HER PARENT who says 'Go ahead, answer it.' The caller is

a professor of DISCIPLINE saying 'Good luck my CHILD, you've got this.' The CHILD ends the call and looks at HIS/HER PARENT, who is still sitting next to HER/HIM in their car. How is this possible?

We manipulated the riddle in three dimensions (with each participant seeing only one version) to test how these features affected responses. First, and most importantly, the riddle featured a PARENT who could be either a mother or a father, which allowed us to examine which gender was more readily associated with a high academic position in a given discipline. Second, the CHILD could be a daughter or a son. Third, we included information about DISCIPLINE to test whether field context influenced perceptions of the professor's gender. The disciplines were selected to represent a male-dominated, a gender-balanced, and a female-dominated field. Of course, the gender composition of a discipline can be measured in several ways – for example, by the ratio of male to female students or faculty members – and none of these indicators may perfectly capture participants' subjective perceptions of gender distribution across departments. We selected physics, art history, and geography because, according to the National Science Foundation's Survey of Earned Doctorates in 2020, women accounted for 19%, 79.8%, and 51.2% of PhD recipients in these fields, respectively. These figures also broadly align with data on bachelor's and/or master's degrees in the United States and the United Kingdom in physics, geography, and art-related fields (Plaister, 2021; Porter & Ivie, 2019; Revell & Benfield, 2022).

Altogether, we used a 2 (PARENT: mother vs. father) × 2 (CHILD: daughter vs. son) × 3 (DISCIPLINE: physics vs. geography vs. art history) between-subject design.

To investigate how implicit gender-science associations correlate with willingness to support women in science, we introduced participants to the American Association of University Women (AAUW), which is a non-partisan, non-profit organization that strives to remove barriers and biases that stand in the way of gender equality, including equal opportunities in education. The participants were informed at the beginning of the survey that one person would be drawn to receive $50. We asked the participants to decide what percentage of this amount, if they won, they would like to donate to AAUW. We informed the participants that, conditional on being drawn, their decision would be implemented and, naturally, we kept the promise. Including this incentivized donation decision allowed us to add a behavioural measure alongside the declarative one, thereby broadening the scope of the analysis.

Subsequently, we collected demographic data. Again, following Belle et al. (2021), in addition to demographic questions, we asked the participants about their political views (conservative, liberal, or moderate) and whether they considered themselves feminists (yes/no questions). Further questions were asked as proxies for participant's exposure to professionally active women. First, we asked them if their mother was professionally active during their childhood. Second, we asked if most of their academic instructors were men or women. Finally, we asked the participants about their main academic field;

using the National Science Foundation data on the percentage of PhDs obtained by women in 2020 mentioned above we thus had a measure of relative feminization of their majors. As in the previous case, we use this as a proxy for exposure to female academics. All of the demographic questions, as well as an example version of the questionnaire, may be found in Supplement 1.

4. Supplementary analyses and validation

As a further measure of implicit bias, participants completed an IAT; its description and results are reported in Supplement 3. Due to the low response rate and the absence of any meaningful correlation with the dependent variable, these results are not discussed in the main article but remain available in Supplement 2.

We also conducted a small supplementary study to assess face validity, examining whether participants regarded identifying the professor as the mother as the most plausible answer and whether failure to do so was perceived as a consequence of stereotype-consistent reasoning; details are provided in Supplement 4.

IV. HYPOTHESES

We formulate the following hypotheses.

Hypothesis 1: The fraction of responders providing the 'normative' answer will be higher in the male professor condition[3] than in the female professor condition.

Hypothesis 2: In the female professor condition, 'normative' answers to the riddle will be associated with higher donations for AAUW.[4]

Hypothesis 3: In the female professor condition, the fraction of 'normative' answers will be positively related to the fraction of female academics in a given discipline.

Hypothesis 4: Women will be more likely than men to answer 'normatively' in the female professor condition.

Hypothesis 5: Respondents with conservative political views will be more likely to give a non-'normative' answer in the female professor condition and more likely to give a 'normative' answer in the male professor condition than those with moderate or liberal political views.

---

[3] In the preregistration, these treatments were labeled the Calling Father Condition and Calling Mother Condition. For clarity, we refer to them here as the male professor condition and the female professor condition.

[4] We have renumbered Hypotheses 2 and 3 relative to the preregistration document to improve the flow of the results section while discussing the results corresponding to each hypothesis in the same order in which the hypotheses are presented.

Hypothesis 6: In the female professor condition, respondents who consider themselves feminists will be more likely to give a 'normative' answer than those who do not consider themselves as feminists.

Hypothesis 7: In the female professor condition, respondents with more exposure to professionally active women will be more likely to give a 'normative' answer.

Hypothesis 8: Higher IAT bias will be associated with a lower chance of giving a 'normative' answer in the female professor condition.

As specified in the preregistered analysis plan, we test the hypotheses using logistic regressions and tests of proportions.[5] Because the IAT was an additional exploratory measure, the logistic regression that includes this variable is reported in the supporting information (Supplement 2). The IAT results were non-significant across all logistic regression specifications.

V. RESULTS

The prevalence of 'normative' answers by condition is shown in Figure 1 and Table 1. The key observation is that the 'normative' answer (the professor is the father) is very common (64.4–77.9% depending on the discipline and child's gender) in the male professor condition; by contrast, the 'normative' answer (the professor is the mother) is very rare (10.3–26.9%) in the female professor condition, providing strong support for Hypothesis 1. The academic field, in violation of Hypothesis 3, had a weaker and less systematic impact on the answers.

Table 1

Number and percentage of 'normative' answers by PARENT, CHILD and DISCIPLINE

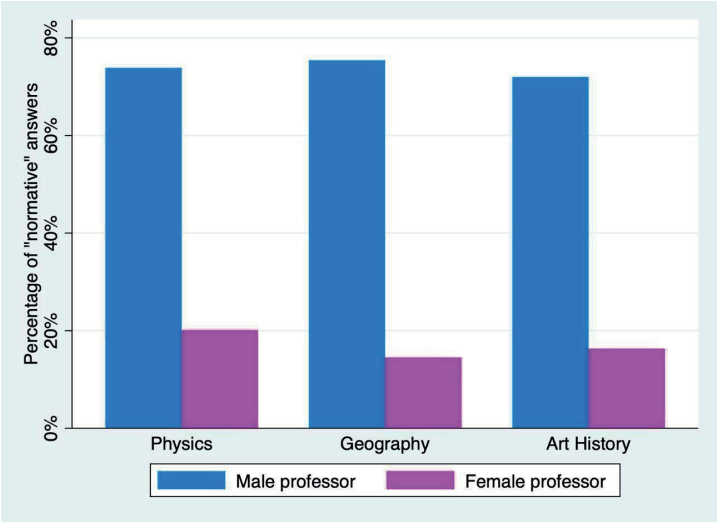
Discipline	Son		Daughter	
	Female prof	Male prof	Female prof	Male prof
Physics	13/96 (13.5%)	84/111 (75.7%)	25/93 (26.9%)	61/87 (70.1%)
Geography	11/107 (10.3%)	81/104 (77.9%)	19/104 (18.3%)	69/95 (72.6%)
Art history	20/124 (16.1%)	92/119 (77.3%)	15/92 (16%)	65/101 (64.4%)

Source: the authors' elaboration prepared using Stata/SE 17 (StataCorp LLC, College Station, TX).

[5] Although we preregistered the Cochran–Armitage trend test, we ultimately relied on logistic regression because the predictor was operationalized as a continuous percentage of feminization in a given discipline, making regression the more appropriate and flexible analytical approach for testing the same directional hypothesis.

Figure 1

Percentage of 'normative' answers by PARENT and DISCIPLINE



Source: the authors' elaboration prepared using Stata/SE 17 (StataCorp LLC, College Station, TX).

In the case of the female professor condition, among the 513 non-'normative' answers, 75.2% referred to the wrong gender. For example, responses such as 'Maybe it's her uncle,' 'One of the fathers is a step-dad,' 'The professor is a priest,' or 'Grandfather called grandson' were common. The most frequent category of non-'normative' answers (31.6%) involved mentioning a same-sex parent, such as 'He has two fathers' or 'The fathers are gay.' Approximately 17.5% of the non-'normative' responses provided various miscellaneous explanations, including 'magic', 'hallucination', 'wrong number', 'voicemail', or 'WTF'. Additionally, 6.6% of the non-'normative' answers indicated that the child is adopted.

On the other hand, in the male professor condition, some participants – typically providing the 'normative' answer – also asked questions such as 'Why is it a riddle?' or commented they were confused ('I'm not entirely sure what the riddle is asking,' 'For that riddle I have no idea what was supposed to be "impossible"'), implying we were asking about something too obvious. Out of the 165 non-'normative' answers given in the male professor condition, only 32.1% indicated the female gender (e.g. 'Lesbian mothers,' 'It was the daughter's grandmother,' 'One of the mothers is a step-mother,' etc.) with only 10.9% participants indicating same-sex marriage, 6.1% indicating adoption 18.2% providing miscellaneous non-'normative' responses, such as 'Lies,' 'She never left the car,' 'Time travel,' 'The daughter is being pranked.'

Some participants (about 3% of the entire sample, i.e. in both PARENT conditions) gave a vaguer answer, for example, saying the professor was 'a parent', 'another parent', without an indication of sex. We categorized these

as non-'normative' answers, as they merely restate the puzzle. However, given their rarity, this categorization decision does not affect the results.

To assess whether recruitment mode may have influenced the results in practice, we compared outcomes between the two groups using a *t*-test. The difference was not statistically significant, $t(1231) = 1.44$, $p = .15$. This provides no evidence that the recruitment strategy (and associated compensation difference) systematically affected the study results.

We now move on to the tests of proportions (always reporting one-sided *p*-values) and logistic regressions. Our dependent variable *normative* is a dummy variable taking the value of 1 if the answer in the female professor condition indicates that it is the mother/woman who is calling or that it is the father/man who is calling in the male professor condition. Otherwise, the variable takes the value of 0.

As can already be inferred from Table 1 and Figure 1 already, the professor's gender had a very large effect on the dependent variable. Aggregating across all disciplines, the prevalence of 'normative' answers was just 16.7% for the female professor condition and 73.3% for the male professor condition ($p = .00$), which confirms Hypothesis 1 (the logistic regression can be found in Supplement 2). The child's gender made a smaller difference, and its direction depended on the professor's gender. In the male professor condition, the prevalence of 'normative' answers was 68.9% for the daughter and 77% for the son ($p = .01$). In the female professor condition, it was 13.5% for the son and 20.4% for the daughter ($p = .02$). The fraction of 'normative' answers was notably higher than could be expected from the main effects in the female professor – daughter – physics condition.

Conversely, the gender of the respondents had a significant impact on the results. In the female professor condition, men were more likely to give a 'normative' answer than women, contradicting our preregistered Hypothesis 4 (21.3% vs. 14%, $p = .01$). However, in the male professor condition, no significant difference was observed between men and women ($p = .96$).

To explore the robustness of these findings, we additionally conducted logistic regressions; see Table 2 for female and male professor conditions. The joint regression for the entire sample can be found in Supplement 2. In every regression, we used three specifications: in Model (1) we included only treatment variables; in Model (2) we also controlled for demographic variables. Finally, in Model (3) we additionally included variables approximating the degree of exposure to a particular gender at the university and donations.

Contrary to Hypothesis 2, the *donation* variable was insignificant across all model specifications. In Hypothesis 3, we assumed that, in the female professor condition, the proportion of 'normative' responses would be positively related to the share of female academics in a given discipline; however, the relationship was negative. Similarly, contrary to Hypothesis 4, women were less likely than men to respond 'normatively' in the female professor condition. Regarding Hypotheses 5 and 6, neither self-identification as a feminist nor declared political orientation showed significant effects.

Table 2

Logistic regression for female and male professor conditions: *'Normative'* as the dependent variable

Variable	Female professor condition			Male professor condition		
	Model			Model		
	(1)	(2)	(3)	(1)	(2)	(3)
Professor's field:						
Geography	0.656 [0.387, 1.112]	0.700 [0.405, 1.207]	0.706 [0.408, 1.224]	1.138 [0.724, 1.789]	1.158 [0.720, 1.864]	1.128 [0.699, 1.822]
Art History	0.793 [0.476, 1.321]	0.833 [0.492, 1.408]	0.835 [0.492, 1.417]	0.918 [0.596, 1.413]	0.938 [0.598, 1.472]	0.936 [0.595, 1.472]
Daughter	1.651** [1.075, 2.536]	1.611** [1.033, 2.512]	1.593** [1.019, 2.492]	0.661** [0.462, 0.947]	0.599*** [0.412, 0.872]	0.595*** [0.408, 0.868]
Respondent's gender:						
Female		0.533*** [0.335, 0.847]	0.628* [0.383, 1.030]		0.798 [0.531, 1.200]	0.908 [0.584, 1.412]
Other		0.325 [0.039, 2.693]	0.328 [0.039, 2.741]		0.421 [0.118, 1.497]	0.483 [0.134, 1.740]
Age		1.009 [0.965, 1.055]	1.015 [0.970, 1.061]		0.972* [0.945, 1.000]	0.975* [0.947, 1.003]
Mother's professional activity		1.005 [0.999, 1.011]	1.005* [0.999, 1.011]		1.001 [0.996, 1.006]	1.001 [0.996, 1.006]
Feminism		1.357 [0.812, 2.267]	1.406 [0.836, 2.366]		1.340 [0.852, 2.108]	1.372 [0.869, 2.166]
Ethnicity:						
Hispanic		1.173 [0.629, 2.189]	1.183 [0.632, 2.214]		0.975 [0.593, 1.602]	0.941 [0.570, 1.556]
Black		0.941 [0.257, 3.454]	0.895 [0.241, 3.322]		0.746 [0.304, 1.832]	0.704 [0.285, 1.741]

Table 2 *(continued)*

Variable	Female professor condition			Male professor condition		
	Model			Model		
	(1)	(2)	(3)	(1)	(2)	(3)
Pacific Islander		(empty)	(empty)		0.557 [0.047, 6.563]	0.589 [0.049, 7.031]
Asian		1.090 [0.596, 1.991]	0.965 [0.519, 1.792]		0.726 [0.440, 1.199]	0.675 [0.405, 1.124]
Other		1.697 [0.732, 3.936]	1.691 [0.723, 3.956]		0.750 [0.278, 2.027]	0.713 [0.261, 1.947]
Political view:						
Liberal		0.809 [0.392, 1.668]	0.835 [0.401, 1.737]		0.599 [0.322, 1.116]	0.624 [0.334, 1.167]
Moderate		1.153 [0.611, 2.177]	1.160 [0.609, 2.211]		0.640 [0.361, 1.133]	0.655 [0.367, 1.167]
Donation			0.999 [0.991, 1.007]			0.996 [0.990, 1.002]
Instructors:						
More women			1.202 [0.655, 2.206]			0.961 [0.565, 1.634]
About the same number			1.028 [0.606, 1.745]			0.857 [0.551, 1.334]
Female rate in the respondent's field			0.974* [0.947, 1.001]			0.983 [0.961, 1.006]
Constant	0.193***	0.133***	0.150***	3.313***	10.359***	13.953***
N	616	593	593	617	592	592
Pseudo R^2	0.014	0.039	0.046	0.008	0.029	0.036

Note. Odds ratios reported; values greater (smaller) than 1 correspond to increased (decreased) likelihood of a 'normative' answer. *$p < .1$; **$p < .05$; ***$p < .01$. 95% confidence intervals reported in brackets.

Source: the authors' elaboration prepared using Stata/SE 17 (StataCorp LLC, College Station, TX).

Contrary to Hypothesis 7, in the female professor condition, respondents with greater exposure to professionally active women were less likely to provide a 'normative' response.

The *daughter* condition was significant in all model specifications: it increased the likelihood of a 'normative' response in the female professor condition and decreased it in the male professor condition. In addition, the mother's professional activity was significant in the female professor condition, increasing the likelihood of a 'normative' answer.

Lastly, older participants performed worse in the male professor condition, being less likely to give the 'normative' response. However, conclusions regarding age effects should be drawn cautiously, as 83.7% of participants were 25 years old or younger, and 92.4% were 30 or younger.

All other variables, including the professor's academic field, were insignificant.

VI. DISCUSSION

This study extends the riddle literature in three main ways. First, it shifts the paradigm from the classic surgeon or specialist setting to academia by focusing on the figure of the professor. Second, it tests whether gender bias varies across disciplines with different gender compositions. Third, it complements the main riddle outcome with additional measures, including donation willingness and an IAT. Across these extensions, the main result is clear: participants showed substantial difficulty in identifying the professor as a woman.

This pattern is broadly consistent with prior riddle studies showing that male-default role schemas are highly resistant to revision. Belle et al. (2021) found that only about 30% of respondents identified the surgeon as the mother, whereas Morehouse et al. (2022) reported even higher failure rates depending on wording. In our academic version, the proportion of participants giving the 'normative' answer in the female professor condition was lower still, at 16.7%. Although comparisons across studies should be made cautiously because of differences in wording and design, this suggests that professorship may be at least as strongly associated with men as the roles examined in earlier riddle paradigms.

A central contribution of the present study is that it tests whether this bias varies across academic disciplines. One might expect the male-professor default to weaken in more feminized fields, such as art history, relative to more male-dominated fields, such as physics. Yet we find little evidence that disciplinary gender composition meaningfully attenuates the stereotype. This suggests that the role of professor may itself be strongly coded as male, independently of the gender balance in a specific discipline. At the same time, this finding should be interpreted cautiously, as participants' subjective beliefs about the gender composition of a field may not correspond closely to objective indicators.

Our results also speak to the role of linguistic cues. Kollmayer et al. (2018) showed that gender-fair wording increased mother-consistent answers in the specialist riddle, and Morehouse et al. (2022) similarly found that wording affected performance in the surgeon riddle. In our study, we observed a significantly higher rate of 'normative' responses in the female professor/female daughter condition and a significantly lower rate in the male professor/female daughter condition. Thus, as in earlier studies, linguistic cues appear to modulate responses. Importantly, however, this effect is modest: even when the wording favours a female reading, the overall rate of 'normative' answers remains low. This suggests that linguistic cues can attenuate, but do not eliminate, the default association between professorship and men.

In one respect, our findings differ from Belle et al. (2021). Whereas they found that women were more likely than men to provide the mother solution, we observe the opposite pattern in our academic context. This result is more in line with Stoeger et al. (2004) and may point to the presence of internalized or auto-stereotypical beliefs. One possible interpretation is that women, like men, may internalize dominant cultural schemas linking academic authority with masculinity. This interpretation is consistent with work suggesting that explicitly egalitarian values do not preclude implicit stereotype-consistent judgments (Rahmani Azad et al., 2023). At the same time, this finding should be interpreted cautiously and warrants replication.

Even among participants who reported having encountered a similar riddle before, a substantial share still failed to identify the professor as the mother. In the female professor condition, 28% of those familiar with a similar riddle nevertheless failed to provide the 'normative' answer. This mirrors the broader pattern reported by Belle et al. (2021) and underscores the persistence of gender schemas: prior familiarity with the logic of the puzzle is often insufficient to override the default male interpretation.

Another notable pattern is the relatively high share of responses invoking same-sex parents rather than identifying the professor as the mother. This may reflect increased social salience of diverse family structures. At the same time, it highlights the persistence of the male-professor schema, as many respondents were more willing to reinterpret the family structure than to infer that the professor was a woman. Some less frequent responses referred to adoptive or stepparents, again suggesting flexibility in family-related reasoning while leaving the gendered occupational schema intact.

We also find that having a professionally active mother predicts responses to the riddle. This may indicate that personal exposure to women in non-traditional or economically active roles increases the accessibility of a female-professor interpretation. At the same time, the mechanism remains unclear. Maternal employment alone does not indicate occupational status, prestige, or whether the occupation itself challenged conventional gender roles. Further research would therefore be needed to determine more precisely what type of exposure matters.

By contrast, several other individual characteristics, including political views and feminist identification, do not appear to predict responses in a ro-

bust way. This result is similar to that of Belle et al. (2021) and may reflect the strength of the underlying schema: when the normative response in the female professor condition is rare, the scope for detecting more subtle correlates is naturally limited. A related consideration may also help explain the lack of significant differences in donation behaviour, which is additionally constrained by a smaller effective sample size for that measure.

The study also makes a methodological contribution. Compared with the classic surgeon riddle, the professor version is less dramatic, less tied to emergency professions, and more readily adaptable to the academic context. This makes it possible to examine gender schemas in a domain where authority, expertise, and leadership are central. It also allows comparisons across disciplines that are more directly comparable to one another than occupations such as surgeon, firefighter, or nurse. In this sense, the present design extends the riddle paradigm to a new institutional setting while preserving its core advantage as an indirect measure of stereotype-consistent inference.

Taken together, the findings suggest that the default association between professorship and men is both strong and surprisingly insensitive to variation in disciplinary gender balance. The most striking result is that only 16.7% of participants in the female professor condition provided the 'normative' answer, compared with 73.3% in the male professor condition. Although direct comparisons should again be made cautiously, this percentage is lower than the corresponding rates reported in several earlier studies using female surgeons, specialists, or doctors (Belle et al., 2021; Kollmayer et al., 2018; Morehouse et al., 2022; Stoeger et al., 2004). This suggests that male-default inferences are not confined to medicine or emergency professions, but also characterize academic authority. At the same time, because our design differs from earlier versions of the riddle and our sample has distinctive features, these conclusions should be interpreted with caution and tested in future replications.

Overall, the results indicate that the figure of the professor remains strongly masculinized in participants' implicit reasoning. Strikingly, this pattern persists even across disciplines that differ substantially in gender composition. The findings therefore extend prior riddle-based research by showing that male-default schemas apply in the academic domain and may remain robust even where objective field characteristics might be expected to weaken them. More broadly, the study suggests that academia continues to operate as a gendered institution not only through formal structures and evaluations, but also through implicit assumptions about who can be recognized as a legitimate academic authority.

VII. LIMITATIONS OF THE STUDY

One limitation of our method is its hypothetical nature. Respondents were asked to imagine a story and solve a hypothetical task. They thus had no incentive to provide the 'normative' answer. That said, it seems unlikely

that such incentives, even if they encouraged participants to think longer and harder, would make a large difference in a riddle of the type considered here.

Another limitation concerns the discipline manipulation. We varied the field associated with the professor (physics, geography, or art history) using gender ratios among PhD recipients as an objective proxy for differences in gender composition across disciplines. However, stereotype activation is likely influenced primarily by participants' perceptions of gender representation within these fields rather than by objective demographic indicators. Such perceptions may reflect a variety of sources, including educational experience or broader cultural narratives about academic disciplines. Moreover, gender ratios among PhD recipients do not necessarily correspond to gender ratios among professors, which may be more directly relevant for judgments about professorial roles. Future research could therefore benefit from directly measuring perceived gender composition across disciplines and validating such manipulations more explicitly.

Another limitation of the study might be the limited reliability of self-report questions, such as 'Are you a feminist?' or 'What are your political views?' While these questions provide valuable insights into participants' self-perceptions, they are subject to biases and subjective interpretations. Using more objective measures or standardized scales could improve reliability; however, implementing such methods was considered too time-consuming and impractical for the scope of our study. Moreover, using a proxy to measure participants' exposure to female researchers may affect reliability, as it is a rough approximation and does not account for many other potential sources of exposure, such as television, social media, or other media platforms.

VIII. CONCLUSIONS

The present study suggests that the professor role remains strongly masculinized in participants' spontaneous reasoning. Even when the riddle clearly allowed for a female professor, only a minority of respondents generated that interpretation, whereas the corresponding male-professor answer was readily available. This asymmetry indicates that academic authority continues to be cognitively associated with men.

This finding matters for research on academia as a gendered institution. Existing scholarship has shown that gender bias shapes evaluations, career trajectories, and access to opportunities within higher education. Our results add a complementary micro-level perspective: they show that gendered institutional patterns are also reflected in basic interpretive expectations about who can plausibly occupy the role of professor. In other words, the masculinization of academic authority appears not only in organizational outcomes, but also in ordinary acts of inference.

The study also refines the riddle paradigm by adapting it to an academic setting and by comparing disciplines that differ in gender composition. The

fact that the stereotype remained strong even in more feminized fields suggests that increasing women's numerical presence alone may not be sufficient to weaken the male-default image of professorship. At the same time, the effects of the daughter condition and maternal professional activity indicate that contextual cues and lived experience can modestly shift responses.

Taken together, the findings point to a persistent gap between social change in academia and the cognitive schemas through which academic roles are perceived. The contribution of the present study lies in showing that this gap can be captured with a simple riddle-based instrument and that the resulting pattern is consistent with viewing academia as a gendered institution reproduced, in part, through everyday stereotype-consistent interpretation.

Such studies demonstrating that certain professions are perceived as predominantly masculine are very important as they can lead to a broader conversation about societal attitudes, potentially influencing policies and initiatives designed to create a more equitable and inclusive environment. Various actions can address this issue. Governments and NGOs can launch awareness campaigns to challenge gender-related biases, aiming to change public attitudes. Companies can implement diversity programs for balanced gender representation, fostering inclusivity in the workplace. Policymakers may invest in educational initiatives to challenge stereotypes from a young age. Educational institutions can adopt policies supporting students irrespective of gender stereotypes. Finally, if biases affect hiring or promotions, policymakers might strengthen legal protections against workplace gender discrimination.

Another potential countermeasure would also be to ensure the fair representation of female academics in the media, as currently, the portrayal of a typical scientist on television often aligns with the image of a white man (Dudo et al., 2011; Long et al., 2010). Actively promoting diverse depictions of scientists, particularly women, in the media, can help challenge and overcome the persistent stereotype of the male professor. Second, increasing the visibility of female scientists could be achieved by providing more information about the gender of authors in academic articles. By explicitly highlighting the gender of authors, such measures create an opportunity to acknowledge and showcase the contributions of female scientists. This can serve as a small but significant step toward increasing the exposure and recognition of women in the scientific community.

Author contributions / Indywidualny wkład autora (CRediT): **Patrycja Janowska-Widomska** – 70% (Conceptualization / Konceptualizacja; Data curation / Zarządzanie danymi; Formal analysis / Formalna analiza; Investigation / Przeprowadzenie badań; Methodology / Metodologia; Resources / Zasoby; Software / Oprogramowanie; Validation / Walidacja; Visualization / Wizualizacja; Writing – original draft / Pisanie – pierwszy szkic; Writing – review & editing / Pisanie – recenzja i edycja). **Michał Krawczyk** – 20% (Conceptualization / Konceptualizacja; Funding acquisition / Pozyskanie funduszy; Investigation / Przeprowadzenie badań; Methodology / Metodologia; Project administration / Administracja projektu; Resources / Zasoby; Supervision / Nadzór; Validation / Walidacja; Writing – original draft / Pisanie – pierwszy szkic; Writing – review & editing / Pisanie – recenzja i edycja). **Dominika Maison** – 10% (Conceptualization /

Konceptualizacja; Methodology / Metodologia; Validation / Walidacja; Writing – review & editing / Pisanie – recenzja i edycja).

Conflict of interest / Konflikt interesów: The authors declare no conflict of interest. / Autorzy nie zgłosili konfliktu interesów.

Funding / Finansowanie: We acknowledge financial support from the National Science Centre (Poland), grant no. 2019/35/O/HS4/01640. / Badanie zostało sfinansowane ze środków Narodowego Centrum Nauki, grant nr 2019/35/O/HS4/01640.

The use of AI tools / Wykorzystanie narzędzi AI: The authors declare no use of AI tools. / Autorzy oświadczyli, że nie korzystali z narzędzi AI.

Data availability / Dostępność danych: The data is available on request. / Dane dostępne na życzenie.

Supplementary materials / Materiały dodatkowe: All materials, anonymized data, and supplementary analyses associated with this study are openly available on the Open Science Framework (OSF, https://osf.io/gu36b/overview?view_only=196df620b412483bbda46991a64cd907). This OSF repository includes all supplementary materials, the fully anonymized dataset, and the relevant documentation. / Wszystkie materiały, zanonimizowane dane oraz analizy uzupełniające związane z tym badaniem są publicznie dostępne w Open Science Framework (OSF, https://osf.io/gu36b/overview?view_only=196df620b412483bbda46991a64cd907). Repozytorium OSF obejmuje wszystkie materiały uzupełniające, w pełni zanonimizowany zbiór danych oraz odpowiednią dokumentację.

Ethics approval statement / Zgoda komisji etycznej: The study has received ethics committee approvals from the Faculty of Economic Sciences, University of Warsaw and Institutional Review Board, Texas A&M University (reference no. 143744). / Badanie uzyskało zgody komisji etycznych Wydziału Nauk Ekonomicznych Uniwersytetu Warszawskiego oraz Institutional Review Board Texas A&M University (nr referencyjny 143744).

Acknowledgement / Podziękowania: We would like to thank Deborah Belle, who provided us with a protocol and questionnaire used in the study by Belle et al. (2021). This allowed for standardization of studies and thus comparability of results. Thanks are also due to Catherine Eckel, who provided valuable guidance and without whom conducting a survey on such a large sample would have been impossible. / Autorzy pragną podziękować Deborah Belle za udostępnienie protokołu badawczego oraz kwestionariusza wykorzystanego w badaniu Belle i in. (2021), co umożliwiło standaryzację badań, a tym samym porównywalność wyników. Podziękowania należą się również Catherine Eckel za cenne wskazówki i wsparcie logistyczne, dzięki którym możliwe było przeprowadzenie badania na tak dużej próbie.

References / Bibliografia

Belle, D., Tartarilla, A. B., Wapman, M., Schlieber, M., & Mercurio, A. E. (2021). "I Can't Operate, That Boy Is My Son!": Gender schemas and a classic riddle. *Sex Roles*, *85*(3–4), 161–171. https://doi.org/10.1007/s11199-020-01211-4

Bem, S. L. (1981). Gender schema theory: A cognitive account of sex typing. *Psychological Review*, *88*(4), 354–364.

Boring, A. (2017). Gender biases in student evaluations of teaching. *Journal of Public Economics*, *145*, 27–41. https://doi.org/10.1016/j.jpubeco.2016.11.006

Carli, L. L., Alawa, L., Lee, Y., Zhao, B., & Kim, E. (2016). Stereotypes about gender and science: Women ≠ scientists. *Psychology of Women Quarterly*, *40*(2), 244–260.

Di Tullio, I. (2019). Gender equality in STEM: Exploring self-efficacy through gender awareness. *Italian Journal of Sociology of Education*, *11*(3), 226–245.

Dudo, A., Brossard, D., Shanahan, J., Scheufele, D. A., Morgan, M., & Signorielli, N. (2011). Science on television in the 21st century: Recent trends in portrayals and their contributions to public attitudes toward science. *Communication Research*, *38*(6), 754–777.

Eagly, A. H., & Diekman, A. B. (2005). What is the problem? Prejudice as an attitude-in-context. In J. F. Dovidio, P. Glick, & L. A. Rudman (Eds.), *On the nature of prejudice: Fifty years after Allport* (pp. 19–35). Blackwell Publishing. https://doi.org/10.1002/9780470773963.ch2

Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review*, *109*(3), 573–598.

Fan, Y., Shepherd, L. J., Slavich, E., Waters, D., Stone, M., Abel, R., & Johnston, E. L. (2019). Gender and cultural bias in student evaluations: Why representation matters. *PLOS ONE*, *14*(2), e0209749. https://doi.org/10.1371/journal.pone.0209749

Fung, J., Fung, W., Israel Rosales, A., Jin, J., & Pettit, R. M. (2023). The role of implicit biases and explicit attitudes toward the poor in donation choices. *Nonprofit and Voluntary Sector Quarterly*, *52*(1), 153–175.

Kollmayer, M., Pfaffel, A., Schober, B., & Brandt, L. (2018). Breaking away from the male stereotype of a specialist: Gendered language affects performance in a thinking task. *Frontiers in Psychology*, *9*, Article 985. https://doi.org/10.3389/fpsyg.2018.00985

Kugler, A. D., Tinsley, C. H., & Ukhaneva, O. (2021). Choice of majors: Are women really different from men? *Economics of Education Review*, *81*, 102079. https://doi.org/10.1016/j.econedurev.2021.102079

Leslie, S.-J., Cimpian, A., Meyer, M., & Freeland, E. (2015). Expectations of brilliance underlie gender distributions across academic disciplines. *Science*, *347*(6219), 262–265. https://doi.org/10.1126/science.1261375

Long, M., Steinke, J., Applegate, B., Knight Lapinski, M., Johnson, M. J., & Ghosh, S. (2010). Portrayals of male and female scientists in television programs popular among middle school-age children. *Science Communication*, *32*(3), 356–382.

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*(4), 291–303. https://doi.org/10.1007/s10755-014-9313-4

Mastekaasa, A., & Smeby, J.-C. (2008). Educational choice and persistence in male- and female-dominated fields. *Higher Education*, *55*, 189–202. https://doi.org/10.1007/s10734-006-9042-4

Miller, D. I., Nolla, K. M., Eagly, A. H., & Uttal, D. H. (2018). The development of children's gender-science stereotypes: A meta-analysis of 5 decades of U.S. Draw-A-Scientist studies. *Child Development*, *89*(6), 1943–1955. https://doi.org/10.1111/cdev.13039

Morehouse, K. N., Kurdi, B., Hakim, E., & Banaji, M. R. (2022). When a stereotype dumbfounds: Probing the nature of the surgeon = male belief. *Current Research in Ecological and Social Psychology*, *3*, Article 100044. https://doi.org/10.1016/j.cresp.2022.100044

Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., & Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences*, *109*(41), 16474–16479. https://doi.org/10.1073/pnas.1211286109

Okoye, K., Arrona-Palacios, A., Camacho-Zuñiga, C., Hammout, N., Nakamura, E. L., Escamilla, J., & Hosseini, S. (2020). Impact of students' evaluation of teaching: A text analysis of teachers' qualities by gender. *International Journal of Educational Technology in Higher Education*, *17*(1), Article 49. https://doi.org/10.1186/s41239-020-00224-z

Plaister, N. (2021, 17 September). *Which A-Level subjects have the best (and worst) gender balance?* FFT Education Datalab. https://ffteducationdatalab.org.uk/2021/09/which-a-level-subjects-have-the-best-and-worst-gender-balance/

Porter, A. M., & Ivie, R. (2019, March). *Women in physics and astronomy, 2019*. AIP Statistical Research Center. https://aip.brightspotcdn.com/1d/3f/cd3407ebe8798bec8900e80be884/women-in-physics-and-astronomy-2019-1.pdf

Rahmani Azad, Z., Goedderz, A., & Hahn, A. (2023). Self-awareness and stereotypes: Accurate prediction of implicit gender stereotyping. *Personality and Social Psychology Bulletin*, *49*(12), 1695–1708. https://doi.org/10.1177/01461672221120703

Revell, M., & Benfield, M. (2022, December 2). *The state of geography: Data and trends in higher education*. American Association of Geographers. https://doi.org/10.14433/2017.0123

Reynolds, D., Garnham, A., & Oakhill, J. (2006). Evidence of immediate activation of gender information from a social role name. *Quarterly Journal of Experimental Psychology*, *59*(5), 886–903. https://doi.org/10.1080/02724980543000088

Rojek, A. E., Khanna, R., Yim, J. W. L., Gardner, R., Lisker, S., Hauer, K. E., Lucey, C., & Sarkar, U. (2019). Differences in narrative language in evaluations of medical students by gender and under-represented minority status. *Journal of General Internal Medicine*, *34*(5), 684–691. https://doi.org/10.1007/s11606-019-04889-9

Steinpreis, R. E., Anders, K. A., & Ritzke, D. (1999). The impact of gender on the review of the curricula vitae of job applicants and tenure candidates: A national empirical study. *Sex Roles*, *41*(7–8), 509–528. https://doi.org/10.1023/A:1018839203698

Stoeger, H., Ziegler, A., & David, H. (2004). What is a specialist? Effects of the male concept of a successful academic person on the performance in a thinking task. *Psychology Science*, *46*(4), 514–530.