Sex Roles (2023) 88:326–347
https://doi.org/10.1007/s11199-023-01355-z

ORIGINAL ARTICLE

Gender Stereotypes Selectively Affect the Remembering of Highly Valued Professions

Alba Sebastián-Tirado[1] · Sonia Félix-Esbrí[1] · Cristina Forn[1] · Carla Sanchis-Segura[1]

Accepted: 15 February 2023 / Published online: 20 March 2023

Abstract
This study includes two experiments designed to assess the effects of occupational gender-related stereotypes on information processing and memory performance. These two experiments were conducted in two separate cohorts of undergraduate students ($N = 107$ and $N = 96$, respectively). In each of them, we assessed (and confirmed) the presence of an implicit association preferentially linking high status attributes to men using the Implicit Association Test (IAT). We also assessed the effective incorporation of this association into gender-schemata and its consequences for information processing with a memory task that involved remembering the feminine and masculine forms of high or low status professional occupations. Results indicated that, independently of their gender, participants were more likely to forget and less likely to falsely recall the feminine forms of high status professions, whereas the opposite was true for the masculine forms of high status professions. The magnitude of these memory biases was correlated with the IAT scores. Moreover, in agreement with the predictions of gender-schemata theory, these memory biases (and their correlations with IAT scores) were predominantly observed when participants were not adverted that their recall would be evaluated later on (incidental-encoding memory task; Experiment 1), but less so when participants were explicitly instructed to memorize the same feminine and masculine forms of high or low status professional occupations (intentional encoding memory task; Experiment 2). Taken together, these results call into question the notion that gender stereotypes about professional occupations are declining, and they highlight a "men-high-status" association as a major component of these occupational stereotypes.

Keywords Gender stereotypes · Memory bias · Social status · Gender schemata theory · Occupational status · Masculine · Feminine · IAT · Incidental encoding

Gender roles refer to the expectations and beliefs about what appearance and behaviors are socially appropriate for different gender groups. As culturally specific gender roles are learned over time and internalized, individuals come to identify and associate everyday objects, people, and activities as masculine or feminine (Bem, 1981; Martin, 2022; Matheus & Quinn, 2017). Individuals also learn to automatically invoke these gender-specific associations when evaluating and incorporating new information. In other words, they learn to process information and respond to new contingencies through gender schemata that are developed early and continuously reinforced or updated through socialization (Bem, 1981). Therefore, across the entire life-span, gender is invoked, often arbitrarily and unnecessarily, to perceive most aspects of the world, and this gender-schematic processing acts as a deforming lens that results in numerous and varied stereotyped associations, beliefs, preferences, and behaviors that ultimately subserve gender typing, polarization, and inequality (Bem, 1993). In the current study, we investigated whether the gender schemata of young adults include associations between gender and social status that could prompt them to process information about professional occupations in a biased way.

Cristina Forn and Carla Sanchis-Segura are contributed equally to this study.

✉ Carla Sanchis-Segura
csanchis@uji.es

[1] Departament de Psicologia Bàsica, Universitat Jaume I, Clínica I PsicobiologiaSos Baynat, 12071 AvdaCastelló, SN, Spain

Gender Schemata: Operations and Manifestations

Gender schemata are dynamic mental structures that help people organize new information according to gender categories and see the world in gender terms (Bem, 1981). More specifically, as Bem (1983) defined the construct, "*A schema is a cognitive structure, a network of associations that organizes and guides an individual's perception […] in which the interaction between incoming information and an individual's preexisting schema determines what is perceived*" (p. 603–604). Gender-schematic processing involves a spontaneous readiness to sort information about persons, but also about all sort of attributes (e.g., kindness, strength), objects (e.g., clothing, tools) and behaviors/ activities (e.g., sports, professions) into masculine and feminine categories based on what is considered typical and/ or appropriate for men and women within a culture. Gender-schematicity may result in explicit beliefs (e.g., "nursing is for women, not for men") but also in automatic cognitive biases that prioritize schema-congruent information and ignore or oppose schema-incongruent information (e.g., assuming that a man sitting in a nurse's station *must be* a doctor, not a nurse). Thus, gender-schematic people tend to judge themselves/ others, perceive their world, establish their preferences, and regulate their behavior based on gender, whereas gender-aschematic people rarely use gender to define themselves/ other persons or their preferences, organize incoming information, or regulate their behavior.

Gender schemata begin to develop in early childhood when infants start to identify themselves as girls or boys and actively construct categories and mental representations of what men and women do and don't do, how they are defined and valued by others, and how they interact with each other (and with inanimate objects) or respond to different situations (Tobin et al., 2010). Children learn which attributes are linked to their own gender and acquire descriptive and prescriptive rules that guide their own and others' views of themselves, and prompt them to match their behaviors to those they believe are appropriate for their gender, thus keeping them aligned with the norms of their gender group and culture (Bem, 1981; Tobin et al., 2010).

Notably, gender schemata bias cognitive processes by promoting selective attention and preferential processing, valuing, and remembering information about the individual's own sex/gender category and/ or information that is congruent with acquired gender stereotypes (e.g., Bem, 1981; Liben & Signorella, 1980; Martin et al., 2002). Thus, gender schemata and stereotypes are readily activated in language processing, and they slow down and/or deteriorate the processing of gender-incongruent information (e.g., Carreiras et al., 1996; Oakhill et al., 2005; Pyykkönen et al., 2010). Gender schemata tend to facilitate memory for gender-congruent information, but they may also lead to the selective forgetting or distortion of memories for gender-incongruent information, so that incoming environmental inputs become consistent with the individual's pre-existing schemata (e.g., Lenton et al., 2001; Leung, 2020; Stangor, 1988).

In this regard, previous studies have shown that individuals have better memory for gender stereotype-consistent materials than for gender stereotype-inconsistent materials. Thus, for example, Liben and Signorella (1980) observed that boys and girls recall pictures illustrating traditional gender-occupation scenarios (that is, those where the sex of the person performing an activity or occupation matches the one that could be predicted by traditional gender roles) better than pictures portraying non-traditional gender-occupation situations or gender-neutral activities/occupations (pictures where men or women perform neutral behaviors, such as singing). These effects have been replicated and extended by other studies assessing how gender schemata and gender stereotypes bias memory processes in children (e.g., Carter & Levy, 1988; Frawley, 2008; Levy, 1989; Liben & Signorella, 1993; Signorella & Liben, 1984; Stangor & Ruble, 1989; for a meta-analysis, see Signorella et al., 1997), but also in adolescents and adults (e.g., Bem, 1981; Cherney, 2005; Herrmann et al., 1992; Renn & Calvert, 1993; Shapiro, 2009).

Gender Stereotypes and Occupational Segregation

Gender stereotypes stem from an awareness of the discrepant distribution of men and women into social roles in specific contexts, such as work (Koenig & Eagly, 2014). Therefore, it is not surprising that, already from early childhood and even more so later in life (Liben et al., 2001; Solbes-Canales et al., 2020), people perceive some professional occupations as "masculine" or "feminine" (Atli, 2017; Matheus & Quinn, 2017; Treleaven, 2015; White & White, 2006) and express deeply ingrained gender-related stereotypes and attitudes about what men and women are naturally "good at" or "better suited for" (Beggs & Doolittle, 1993; Guo et al., 2015; Halper et al., 2019; Shinar, 1975; Treleaven, 2015). Thus, because men are ordinarily considered competitive, independent, and brave, they are perceived as more appropriate for managerial, challenging, and stressful jobs (e.g., lawyers, surgeons, or executive managers), whereas women are seen as compassionate, delicate and, therefore, better suited to clerical and care-related occupations (e.g., secretaries, teachers, or nurses). These gender stereotypes weigh heavily in career choices (Clarke, 2020; Heilman, 2012), so boys and girls usually aspire -and, as adults, tend to occupy- jobs that are traditionally stereotyped for their gender (Liben et al., 2001;

Sellers et al., 1999; Weisgram et al., 2010), hence promoting and reinforcing pre-existing gender-based occupational segregation and, thereby, perpetuating these same stereotypes.

Because male-typed careers tend to offer more social prestige and authority and are better paid than female-typed careers (Clarke, 2020; García-Ael et al., 2018; Hegewisch et al., 2010), men and women often have unequal levels of power and status in their communities. Gender-based power/status asymmetries in work-related contexts are further reinforced by vertical segregation. Vertical segregation refers to the differential concentration of men and women in different levels of responsibility or authority positions (Salanauskaite, 2017). This segregation is sustained less by the previously mentioned essentialist beliefs about the different skills and temperamental characteristics of men and women than by the generalized (although rarely stated explicitly) assumption that men are less interested in/accountable for domestic responsibilities and more agentic, competent, and committed to work than women (Eagly & Mladinic, 1994; Levanon & Grusky, 2016).This means that men are tracked into better paid and more prestigious careers (Blair-Loy, 2009; Levanon & Grusky, 2016; Ridgeway, 2011) and are privileged by employers in the competition for high-status occupations and positions (Correll et al., 2014; Cotter et al., 2001). Consequently, women are relegated to lesser positions and may even select secondary professional tracks, invest less in their careers and more in their families, or settle for less desirable positions (Blair-Loy, 2009; Charles & Grusky, 2005). Thus, by its mere continuation, vertical segregation reinforces the very same beliefs that prompted it, hence forming a self-perpetuating loop of gender inequality in the workplace, but also, in the domestic domain.

Does Occupational Segregation Result in Gender-Biased Cognitions About Power and Social Status?

In addition to perpetuating pre-existing gender stereotypes, vertical and horizontal segregation provide contexts and create conditions that could lead to the emergence of biased cognitions differentially linking women and men to economic retribution, prestige, and/or authority (i.e., high-status attributes). More specifically, and in accordance with the principles of the social role theory (Koenig & Eagly, 2014) and of the status construction theory (Ridgeway & Erickson, 2000), we propose that the repeated observation of gender-based power/status asymmetries may lead to the development and interiorization of an association directly linking men to high-status attributes and women to lower status attributes. To our knowledge, the existence of this association has not been directly proven. However, several sources of empirical evidence seem to suggest that this association exists and that it is integrated into gender schemata, hence biasing the processing of new and previously stored information.

Thus, children, adolescents, and adults rate male-typed occupations as affording money- and power-related values (Weisgram et al., 2010) and as being higher in status (Liben et al., 2001) than female-typed occupations. However, the lower wage and social status assigned to female-typed jobs seem to stem less from the characteristics of these jobs than from the fact that they are ordinarily performed by women. That is, lower status seems to be primarily ascribed to women and then transferred to the jobs they typically perform. In fact, men and women estimate higher salaries for "masculine jobs" than for "feminine jobs," even when the duties of these jobs are exactly the same (Alksnis et al., 2008), women and men pay themselves less for completing a task labeled as "feminine" than when the same task is labeled as "masculine" (Major & Forcey, 1985), and even novel or non-existing jobs are rated as affording less social status when they are pictured as performed by women (Liben et al., 2001; Weisgram et al., 2010).

On the other hand, women report *deserving* lower incomes than men in past and future jobs (Desmarais & Curtis, 1997), and show a lower sense of personal entitlement to pay and other status-related outcomes (Ciani et al., 2008; Major, 1994), suggesting that the proposed "gender-status" association is deeply interiorized and integrated into schemata that bias self- and others' perception. In fact, there is a large corpus of studies showing that women expect, demand, and allocate lower salaries to themselves than to men, whereas men do the exact opposite (e.g., Callahan-Levy & Messé, 1979; Major et al., 1984; O'Brien et al., 2012; Pelham & Hetts, 2001; Streilein et al., 2018). Moreover, when given the opportunity to pay themselves, women calibrate their lower self-payments based on the task difficulty and their perceived performance, whereas men's higher self-payments are unrelated to these two factors but strongly related to their feelings of self-worth (Pelham & Hetts, 2001).

Therefore, when taken together, these observations seem to suggest that high-status attributes (e.g., high salary) are preferentially linked to men and/ or that low-status attributes are linked to women. The interiorization of this bias and its integration in higher-order cognitive schemata could, together with other factors (e.g., the gendered division of family-related responsibilities; Lips & Lawson, 2009), help to explain several phenomena deriving from an asymmetrical valuation of women and men in the workplace, including why women express higher satisfaction despite being aware that they receive lower salaries than men ("the paradox of the content female worker"; Crosby, 1982), and why girls are interested in female-typed jobs despite rating them as lower in social prestige and salary (Liben et al., 2001). However, and despite its potential relevance, studies assessing whether people directly associate gender and social status and/or how this association is incorporated into gender schemata are lacking.

The Present Study

The first purpose of this study was to explore whether there is a stereotyped association between gender and status held by young, educated people with no direct work experience. To assess whether this proposed implicit stereotyped association actually exists, we employed the Implicit Association Test (IAT; Greenwald et al., 1998). The IAT is a flexible and robust reaction time-based procedure that quantitatively assesses the strength of existing associations between pairs of semantic or social categories at the implicit level (Greenwald et al., 1998; Schnabel et al., 2008). The use of such an implicit measurement is crucial when measuring people's thinking on socially sensitive or potentially self-demoting contents because, unlike explicit measures such as self-reports, implicit measures rely neither on participants' awareness, willingness, nor ability to accurately report these kinds of information (Greenwald et al., 2002).

Given these advantages and its intrinsic methodological flexibility, the IAT has been widely used to assess several kinds of gender-related implicit associations (Nosek et al., 2002, 2007), including the preferentially linking of men to professional career and women to family (Nosek et al., 2002; Salles et al., 2019), linking gender attributes and male/ female typed occupations (Matheus & Quinn, 2017; White & White, 2006), and many others (Dasgupta & Asgari, 2004; Marini & Banaji, 2022; Nosek et al., 2009). Therefore, in the present study, we adapted the general IAT procedure to construct a "gender-status" IAT specifically assessing the here proposed gender-status association, whereas the gender-career IAT (Nosek et al., 2002; Salles et al., 2019) was used to measure the related, albeit distinct, stereotyped association linking men with the public/ professional sphere and women with the domestic/ familiar one.

Gender-related implicit associations become especially relevant when they are incorporated into gender schemata and bias the processing of new and/ or previously stored information. Therefore, the main objective of the present study was to assess whether the gender-status association is incorporated into gender schemata and how this might affect mnemonic processes. To do so, we designed a memory task in which omission (recognition failures) and commission (false memories creation) errors could reveal underlying gender-related biases. In the construction of this task, we took advantage of the fact that, in Spanish, all the nouns (including those referring to professional occupations) are gendered (i.e., the masculine forms of these nouns end with an "o", and the feminine forms end with an "a"), thus making it possible to assess whether presenting the masculine or feminine forms of nouns referring to professional occupations with high/ low status resulted in different levels of correct retrieval. More specifically, our initial hypothesis was that, if the proposed implicit gender-status association is part of a gender scheme that can be activated by gender cues (in this case, grammatical gender markers), the feminine forms of high-status professions should be perceived as gender-incongruent information and, therefore, remembered less well than the masculine form of the same nouns. Finally, because previous studies have shown that the mnemonic manifestations of gender schemata are enhanced when information is encoded without knowing that it should be remembered (e.g., Cherney, 2005), we conducted two separate studies (involving two distinct cohorts of participants) to evaluate the presence and strength of this expected memory bias after explicitly instructing (intentional encoding condition) or not instructing (incidental encoding condition) the participants that their memory of these stimuli would be tested later on.

Hypotheses

Based on all these considerations, four a priori hypotheses were formulated:

H1: Participants will harbor implicit associations preferentially linking men to high-status and professional careers and women to lower status and family, that is, they will show scores larger than zero on the gender-status and gender-career IATs. The strength of these associations might be different in men and women, but no specific predictions are formulated in this regard. However, we do expect these two implicit associations to be positively correlated.

H2: Incidental encoding will result in a low number of correct responses on the recognition test. Achieved performance will be independent of the participant's gender, but the number and type of errors (recognition failures and false recognition memories) will vary for each stimulus category. More specifically, if H1 is confirmed, the frequency of recognition failures (false negative errors; F-) is expected to be higher for the feminine forms of high-status professions (F-HS) than for the masculine forms of high-status professions (M-HS). In addition, the number of false recognition memories (false positive errors; F+) is expected to be higher for M-HS than for F-HS stimuli. The magnitude of these memory biases might be different in women and men, but no specific predictions are formulated in this regard.

H3: Intentional encoding will result in a larger number of correct responses on the recognition test, and in a more balanced distribution of the number and type of errors across the four stimulus categories than that observed after incidental encoding. That is, intentional encoding is expected to attenuate the memory biases described in H2.

H4: The scores of the gender-status and gender-career IATs will be correlated with the memory biases observed in the incidental-encoding memory task described in H2 (but not with any index of memory performance). These correlations are expected to be attenuated or fully suppressed in the intentional-encoding condition (Experiment 2).

Experiment 1

Method

The study was approved by the Ethics Standards Committees of the Universitat Jaume I.

Participants

Participants were 109 undergraduate students from the Universitat Jaume I (Spain) who volunteered in response to an invitational email. The recruited sample had a similar proportion of self-reported men ($n = 52$; 44.7%) and women ($n = 57$; 52.3%) who did not significantly differ in age ($M_{men} = 23.13$, $SD_{men} = 4.379$; $M_{women} = 22.68$, $SD_{women} = 3.031$; $t(107) = 0.629$, $p = .531$). All the participants signed informed consents, and their collaboration was rewarded with €20.

The experiment was carried out in six different experimental sessions, and each session involved 15–20 participants with similar proportions of women and men. At the beginning of each session, three experimenters greeted the participants in the laboratory and randomly assigned them to an individual desk equipped with a personal computer. After the participants had given their informed consent, the experimenters asked them to fill in a demographic data form and wait for further instructions.

Measures and Materials

All the experimental tasks were programmed and presented using the Millisecond Inquisit software package 4.0 (Millisecond©). All responses were automatically collected using the same software. To ensure that all participants understood each experimental task, a lead researcher used a video projection system to provide the pertinent instructions, which were later also displayed on the participants' computer monitors. The tasks and measures included in this experiment were presented in the order listed below.

Incidental-Encoding Memory Task

This task was designed to promote a faint and unintentional encoding of a series of word stimuli (see below) whose recognition was tested 15 min later. Thus, participants were not told that they were performing a memory task until the recognition test took place. In this way, this task aims to maximize the interference of possible participant biases in the recall of the previously presented stimuli.

Forty words were used as stimuli in this task. These words denoted the Spanish masculine ("M") and feminine ("F") grammatical forms of 20 professions, 10 of which referred to highly socially valued occupations (hereinafter labelled "high-status" or "HS"), whereas the other 10 referred to less socially valued occupations (hereinafter labelled "low-status" or "LS"), resulting in four different stimulus categories (F-HS, M-HS, F-LS, and M-LS). These stimuli were selected to ensure that: 1) HS and LS occupations had a very different social value; 2) HS and LS stimuli words had similar lengths; 3) The masculine and feminine form of each word only differed on the last letter.

From the stimulus pool, 20 stimuli (5 of each of the previously defined four categories) were selected as targets and presented during the incidental encoding phase (see below). The remaining 20 stimuli were only presented during the recognition test.

This task was composed of two separate phases: incidental encoding (with 2 incidental encoding runs) and the recognition test.

Incidental Encoding Phase

The task was described to participants as measuring their "perceptual speed" abilities. In the first run, participants were asked to identify (as fast and as accurately as possible) whether the first letter of each of the 20 words individually presented at the center of their computer screens was a vowel (by pressing the "v" key on their keyboard) or a consonant (by pressing the "c" key). The participants' responses were automatically collected, but regardless of the participants' actions, each stimulus remained on the screen for 3 s. The inter-stimuli interval was 0.5 s, and stimuli were presented in a pseudo-random order that avoided displaying more than two items from the same category consecutively. Immediately after completing this first experimental run, participants were told that their "perceptual speed" skills would be re-assessed, but in this case, they had to indicate whether the penultimate letter was a vowel or a consonant. The display of the stimuli and the collection of participants' responses were identical to the previous run.

Recognition Test Phase

The recognition test was conducted 15 min after the incidental encoding phase. During this interim period, participants were asked to perform an unrelated distracting task (a 3D mental rotation task). Once the distracting task was over, participants were told that their mnemonic abilities were going to be evaluated. Thus, participants were asked to declare whether each of the upcoming word stimuli had appeared during the "perceptual speed task." Participants were instructed to give their responses within 3 s by pressing the "b" key (masked with a green tag) on their keyboard if the word had been presented during the "perceptual speed task" or the "n" key (masked with a red tag) if the word had not been presented previously. Immediately after receiving these instructions, all 40 word-stimuli were individually and successively displayed in a random order. Each stimulus was presented at the center of the screen for a maximum of 3 s (inter-stimuli interval: 0.5 s), and the responses provided (yes/ no), or their absence (omissions), were automatically collected. Thus, because 40 stimuli were presented and participants had to choose between two response options, the maximum possible number of correct responses was 40, whereas the number of correct responses expected by chance performance was 20.

Implicit Association Test (IAT)

The Implicit Association Test (Greenwald et al., 1998) is built upon the observation that responding to information perceived as congruent is faster than responding to information perceived as incongruent. More specifically, the IAT measures the strength of the association between concepts and attributes by calculating the standardized reaction time difference (the so-called D scores) between stereotypical non-stereotypical concept-attribute pairs, and it has been widely used to assess several implicit stereotypic associations (Nosek et al., 2002, 2007). In this study, two different versions of this procedure were sequentially implemented: the gender-career IAT and the gender-status IAT, with the latter developed for this study. These tests were administered to all participants just after finishing the memory task.

The gender-career IAT has been in use for more than 15 years, and it provides a measure of the traditional stereotypic association of men with the public/ professional sphere and women with the domestic/ familiar one by specifically testing whether participants preferentially link men to professional career descriptors and women to family-related words (Nosek et al., 2002). Therefore, we employed this script to measure a very general but well-established gender bias, and its scores provide a valid reference to gauge the results obtained with the gender-status IAT. Thus, for this experiment, the freely available IAT gender-career script (the Milisecond Test Library, http://www.millisecond.com/download/library/) was adapted and translated into Spanish (see details in Supplement A in the online supplement). The same computer script was also used as a template when constructing the gender-status IAT. More specifically, the original attribute lists of the gender-career IAT were replaced by new ones containing a series of words referring to three components (social prestige, authority, and economic retribution) of social status, hence providing a measure of the expected stereotypical associations between gender and status (i.e., "men-high-status"/ "women-low status"). The high-status and low-status attributes used to construct the gender-status IAT and the gender-career IAT are provided in Supplement A in the online supplement.

Analytic Strategy

Performed analyses included exploratory and confirmatory methods. Although directional a priori hypotheses were established, all the employed tests were two-tailed, so statistical significance for effects aligned with a priori predictions was not favored over alternative or unanticipated effects. Moreover, following current recommendations (Wasserstein & Lazar, 2016), analyses were not restricted to statistical significance testing and gave equal importance to effect size estimation.

An a priori power analysis was conducted using G*Power3 (Faul et al., 2007) to estimate the required sample size when testing the difference between two independent group means using a two-tailed Student's *t*-test, assuming a medium effect size ($d = .60$), and an alpha of .05. Result showed that a total sample of 90 participants (two equal-sized groups of $n = 45$) was required to achieve a power of .80. A posteriori sensitivity analyses assuming the same parameters and attending to the achieved sample sizes in experiments 1 and 2 were also conducted. The results of these tests indicated that mean differences with $d \geq 0.54$ and $d \geq 0.57$ could be reliably detected in experiments 1 and 2, respectively.

The improved scoring algorithm recommended by Greenwald et al. (2003) was used to calculate the D scores for each participant's IAT responses and as conventionally done, positive D scores denote stereotypical implicit associations (e.g., "women-family", "men-career") whereas negative D scores denote counter-stereotypical (e.g., "women-career", "men-family") associations (see Supplement A in the online supplement for details about the calculation of these scores). The obtained D scores were subsequently analyzed by means of Students' *t*-tests. Thus, one-sample *t*-tests were used to ascertain whether the men's and women's D score averages differed from zero, whereas *t*-tests for independent samples were used to compare these averages between genders. When a statistically significant effect was found in any of these *t*-tests, Cohen's *d* and its corresponding 95% confidence intervals were calculated to estimate the size of the observed effect.

From the incidental memory task, two distinct sets of dependent variables were calculated. First, the averages of correct responses, errors, and omissions provided adequate information about the participants' memory performance and allowed us to investigate possible differences between women and men through a MANOVA followed by univariate ANOVAs. Additionally, a one-sample *t*-test was used to assess whether the average number of correct responses was different from what would be expected by chance (20 correct responses), and the size of this effect was quantified in terms of Cohen's *d* and the corresponding 95% confidence intervals and as the percentage of the maximum possible score (Cohen et al., 1999).

In addition, the averages of false negative errors (F-) and false positive errors (F+) were compared using a 2 (participants' gender: men vs. women) × 2 (grammatical gender: M vs. F) × 2 (stimuli social value: HS vs. LS) ANOVAs. Appropriate effect size indices (eta-squared) were calculated for each statistically significant main or interaction effect. Specific mean comparisons were conducted using paired *t*-tests (with significance levels properly corrected for multiple comparisons using the Bonferroni method), whereas the size of the observed effect was quantified in terms of Cohen's *d* and the corresponding 95% confidence intervals. These statistical analyses were conducted employed the functions included in the *rstatix* package for R.

These subject-based comparisons were complemented by stimuli-focused analyses. Thus, the Chi-square goodness of fit test was used to determine whether the accumulated frequencies of F- and F+ errors were evenly distributed among the four stimulus categories. When a statistically significant effect was found, the standardized residuals (whose interpretation is analogous to that of Z scores; Agresti, 2007) allowed us to identify the cells that deviated significantly from expected values. Moreover, relative risks (RR) and their 95% confidence intervals were calculated to provide direct information about the differential likelihood of F- and F+ errors in each stimulus category. These statistical analyses were conducted employed functions included in the *base* and *epitools* packages for R.

Finally, Spearman's rho correlation index was used to assess whether the frequencies of F- and F+ errors for the different stimulus categories formed a reciprocal and coherent pattern and ascertain their possible relationships with the IAT scores. These correlations were calculated with the *Hmisc* package for R.

Results

Implicit Gender-Career and Gender-Status Associations

H1 anticipated that participants would display implicit associations preferentially linking men to high-status attributes and women to low-status attributes. This hypothesis was tested using two separate one-sample Student's *t*-tests that assessed whether the men and women mean scores in the gender-status IAT (see Table 1 and also Fig. S1 in the in the online supplement) were significantly larger than zero (the "no-bias" value). The obtained results confirmed the presence of this bias in men, $t(51) = 8.51$, $p < .001$, $d = 1.18$, 95% CI [0.88, 1.59], as well as in women, $t(56) = 6.48$, $p < .001$, $d = 0.86$, 95% CI [0.61, 1.21]. On the other hand, to assess whether the strength of this bias differed between women and men, their IAT scores were compared by employing a Student's *t*-test for independent samples. This comparison did not reach statistical significance, $t(107) = 1.09$, $p = .276$, $d = 0.21$, 95%CI [-0.18, 0.57], indicating that the magnitude of this bias was similar in men and women.

H1 also anticipated that both men and women would implicitly associate "men" with "professional career" and "women" with "family." To test this hypothesis, two separate one-sample Student's *t*-tests examined whether men's and women's mean scores in the gender-career IAT were significantly larger than zero (see Table 1 and Supplementary Fig. S1). The obtained results confirmed that men, $t(51) = 7.86$, $p < .001$, $d = 1.09$, 95% CI [0.76, 1.58], and women, $t(56) = 8.97$, $p < .001$, $d = 1.19$, 95% CI [0.83, 1.76], exhibited this cognitive bias. However, a Student's *t*-test for independent samples revealed that the strength of this implicit association did not significantly differ between women and men, $t(107) = 0.06$, $p = .954$, $d = -0.01$, 95% CI [-0.38, 0.35].

Finally, H1 predicted that the observed D scores for the gender-status and gender-career IATs would be related to each other. This prediction was assessed using Spearman's rank-order correlation index, which confirmed a statistically significant association between these two cognitive biases, rho $(107) = .35$, $p < .001$.

Memory Performance and Memory Biases

Memory Performance

H2 stated that incidental encoding would result in poor performance on the recognition test. This hypothesis was tested by using a one-sample Student's *t*-test that compared the average number of correct responses on the recognition test

Table 1 Descriptive Statistics for the Gender-Status and Gender-Career IAT in Experiment 1

	Men		Women	
	M	(*SD*)	*M*	(*SD*)
IAT gender-status	0.31	(0.27)	0.26	(0.3)
IAT gender-career	0.27	(0.25)	0.27	(0.23)

M mean, *(SD)* standard deviation

exhibited by all participants (21.90) to the number of correct responses that could be expected by chance (20). The obtained results confirmed that the observed performance was significantly different but just slightly higher than what would be expected by chance, $t(108) = 4.62$, $p < .001$, $d = 0.44$, 95% CI [0.23, 0.72]. Expressed in more meaningful terms, the participants' performance equated to 54.75% of the maximum possible correct responses in this task, which is barely above the expected random performance level (50%).

H2 also anticipated that performance in the incidental memory task would be independent of the participants' gender. This prediction was supported by the results of a one-way MANOVA, Wilks' lambda (2, 106) = 0.99, $p = .966$, which revealed that women and men did not statistically differ on the number of correct responses, omissions, or errors (the outcomes of the associated univariate ANOVAs, as well as the groups' descriptive statistics, are provided in Table 2; see also Fig. S2 in the online supplement).

Of note, in this task participants could incur two types of errors. More specifically, participants could fail to recognize previously presented words (false negative errors, F) but they could also falsely recognize non-presented words (false positive errors, F+). Because these two types of errors are mutually independent, women and men could potentially differ in the relative amount of F- and F+ even if they do not differ in the total number of errors. To assess this possibility, an additional one-way MANOVA comparing the global averages of women and men on F- and F+ errors was conducted. The obtained results, Wilks' lambda (2, 106) = 0.99, $p = .934$, indicated that women and men incurred a similar number of F+ and F- errors (see associated one-way ANOVAs and the groups' descriptive statistics in Table 2).

Memory Biases

A final prediction included in H2 was that the number and type of errors (F+ and F-) would vary for the different stimuli categories. This prediction was first assessed by comparing the averages of these two types of errors with two separate three-way ANOVAs that included the participants' gender (men vs. women), grammatical gender (M vs. F), and stimuli social value (HS v.s LS) as main factors. To complement these subject-based comparisons, stimuli-focused analyses were also performed. More specifically, two separate Chi-squared (goodness of fit) tests were employed to evaluate whether F- and F+ errors were evenly distributed among the four stimulus categories, respectively. The results of all these analyses are described below.

Recognition Failures (F- Errors) for Each Stimulus Category As hypothesized in H2, the average number of forgotten stimuli was independent of the participants' gender but differed across the stimulus categories More specifically, a 2 (participants' gender) × 2 (grammatical gender stimuli) × 2 (social value stimuli) ANOVA yielded a significant effect of the grammatical gender stimuli factor, $F(1,107) = 23.26$, $p < .001$, $\eta 2 = .18$, the social value factor, $F(1,107) = 24.98$, $p < .001$, $\eta 2 = .19$, and their interaction, $F(1,107) = 41.13$, $p < .001$, $\eta 2 = .28$. Conversely, neither the participants' gender factor nor any of the interactions involving this factor reached statistical significance. Therefore, the stimuli's characteristics (grammatical gender and social value), but not the participants' gender, had a statistically significant influence on the number of recognition failures. More specifically, as Fig. 1a shows, the feminine forms of high-status professions were forgotten more often than their masculine forms, $t(108) = 7.84$, $p_{Bonferroni} < .001$, $d = 1.05$, 95% CI [0.64, 1.55]. Of note, this effect of grammatical gender was not found for stimuli referring to less socially valued jobs, $t(108) = 0.17$, $p_{Bonferroni} = 1.00$.

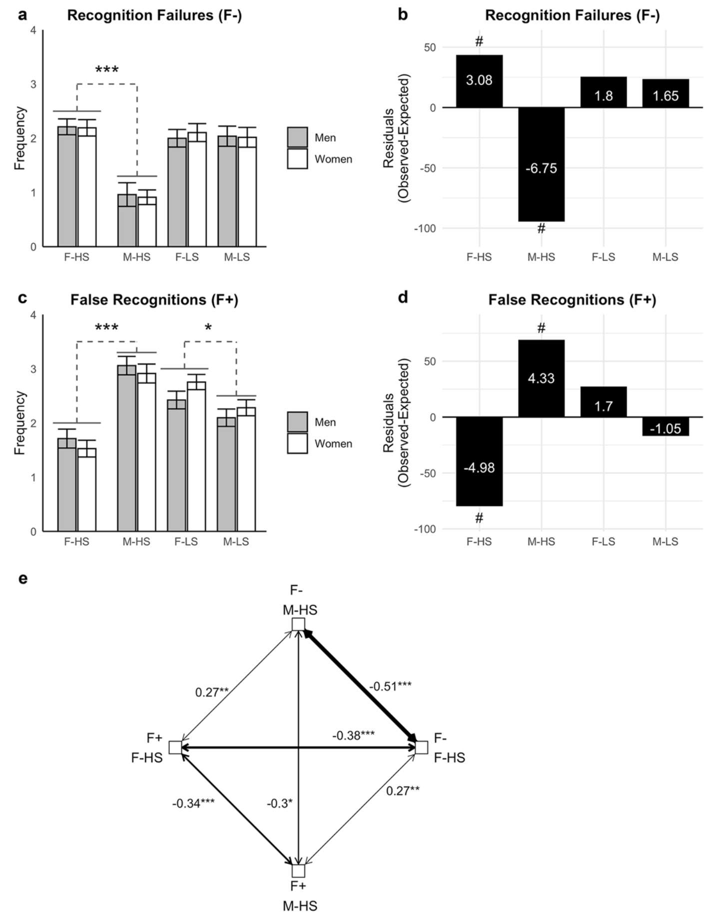
Confirming and extending these results, a chi-square test revealed that F- errors were not evenly distributed among the four word-stimulus categories, $\chi 2\ (3) = 61.90$, $p < .001$. More specifically, and as illustrated in Fig. 1b, participants failed to recognize the feminine forms of high-status professions more often than expected ($Z_{\text{F-HS}} = 3.08$), whereas the opposite was true for the masculine forms of high-status professions ($Z_{\text{M-HS}} = -6.75$). In fact, the probability of failing to recognize the feminine forms of high-status professions was twice as high as that observed for their masculine counterparts, RR = 2.35, 95% CI [1.89, 2.92], $Z = 7.76$, $p < .001$.

Table 2 Descriptive Statistics and Gender-Based Comparisons for Memory Performance Indexes in Experiment 1 (Incidental Encoding)

	Men		Women		Comparison	
	M	*(SD)*	*M*	*(SD)*	*F*(1,107)	*p*
Correct responses	21.92	(4.16)	21.87	(4.43)	0.01	.920
Total errors	16.5	(3.56)	16.7	(4.02)	0.08	.785
False positive errors (F+)	9.28	(2.7)	9.47	(2.6)	0.13	.716
False negative errors (F-)	7.22	(2.89)	7.22	(3.24)	0.001	.978
Omissions	1.56	(2.2)	1.43	(1.26)	0.12	.728

M mean, *(SD)* standard deviation. *F* and *p* values refer to one-way ANOVAs comparing women's and men's scores on memory performance indexes

Fig. 1 Memory Errors and Biases after Incidental Encoding



Note. Panels a and c depict the means ± SEM of recognition failures (false negative errors) and false recognition memories (false positive errors) displayed by women (white) and men (gray) in each stimulus category (F-HS, high-status professions in their feminine grammatical forms; M-HS, high-status professions in their masculine grammatical forms; F-LS, low-status professions in their feminine grammatical forms; M-LS, low-status professions in their masculine grammatical forms), respectively. Panels b and d depict the participants' (men and women) residual scores of false negative and false positive errors for each stimulus category, respectively. The values of the standardized residuals are provided within each bar and, in these two panels, the # symbol is used to denote standardized residuals higher than 2 or smaller than -2 and. therefore, statistically significant departures from the chance-expected number of errors. Finally, Panel e illustrates the pattern of rank order correlations (Spearman's rho) between false negative and false positive errors for the feminine (F-HS) and masculine (M-HS) forms of highly socially valued professional careers. In this panel, the arrows' width denotes the strength of the observed correlations (which exact values are provided within the figure), whereas the color denotes whether these correlations reached statistical significance (black) or not (gray). In all panels, *$p < .05$, **, $p < .01$, *** $p < .001$.

On the other hand, the frequencies of F-errors for stimuli referring to feminine or masculine forms of low-status professions did not deviate from the expected values (Fig. 1b) nor seem to mutually differ, RR = 1.01, 95% CI [0.86, -1.2], $Z = 0.16$, $p = .873$.

False Recognitions (F + Errors) for each Stimulus Category As predicted in H2, the number of non-previously presented items that were erroneously recognized (false recognitions, F+) varied across the different stimulus categories but it did not differ between women and men. More specifically, a 2 (participants' gender) × 2 (grammatical gender stimuli) × 2 (social value stimuli) ANOVA only revealed significant effects for the grammatical gender stimuli, $F(1,107) = 20.42$, $p < .001$, $\eta 2 = .16$, and the grammatical gender x social value interaction, $F(1,107) = 62.79$, $p < .001$, $\eta 2 = .37$. Thus, as depicted in Fig. 1c, participants created fewer false recognition memories for feminine compared to masculine forms of high-status professions, $t(108) = -8.38$, $p_{Bonferroni} < .001$, $d = -1.10$, 95% CI [-1.44, -0.82]. Moreover, F + errors were more frequently observed for the feminine (F-LS) than for the masculine (M-LS) forms of low-status professions, $t(108) = 2.81$, $p_{Bonferroni} < .05$, although this unanticipated effect was much smaller than the effect observed for nouns referring to high-status professions, $d = 0.36$, 95% CI [0.09, 0.65].

These results were extended by a chi-square goodness-of-fit test which confirmed that false recognitions were not randomly distributed across the four stimulus categories, $\chi 2(3) = 47.62$, $p < .001$. More specifically, as shown in Fig. 1d, participants created more false recognition memories for the masculine forms of high-status professions than what would be expected by chance, $Z_{\text{M-HS}} = 4.33$, whereas the opposite was true for the feminine forms of high-status professions, $Z_{\text{F-HS}} = -4.98$. Thus, the probability of committing F + errors on feminine forms of high-status professions was about half as likely as what was observed for masculine forms of high-status professions, RR = 0.54, 95% CI [0.46, 0.64], $Z = 7.31$, $p < .001$. Conversely, neither the number of F + errors for the feminine (F-LS) nor the masculine forms of low-status professions (M-LS) significantly differed from what would be expected by chance; yet, the risk of creating false recognition memories feminine forms of low status professions was just slightly higher than that for the masculine forms of low status professions, RR: 1.18, 95% CI [1.02, 1.38], $Z = 2.2$; $p < .05$.

Relationship Between F- and F + Errors for HS Stimuli Our previous analyses revealed that the frequency of recognition failures (F-) and false recognition memories (F+) on socially valued occupations (HS stimuli) is not random, but rather dependent on the grammatical gender (a dependency that was not observed for less valued occupations; LS stimuli). Moreover, although F- and F + errors on feminine forms of high-status professions or masculine forms of high status professions occur in mutually independent trials, they seem to follow an inverse and complementary pattern. This subjective impression prompted us to quantitatively explore these possible relationships through a full correlation analysis (Fig. 1e).

Specifically, a negative correlation was observed between the number of recognition failures and the number of false recognition memories for the feminine forms of high-status professions, rho(107) = -.38 $p < .001$. A negative correlation was also observed between the number of recognition failures for the feminine forms of high-status professions and the number of recognition failures for masculine forms of high-status professions, rho(107) = -.51, $p < .001$. A positive correlation was observed between the number of recognition failures for the feminine forms of high-status professions and the number of false recognition memories for the masculine forms of high-status professions, rho(107) = .27, $p < .005$. The frequency of these false recognition memories for the masculine forms of high status professions was, in turn, negatively correlated with the frequency of recognition failures for the masculine forms of high status professions, rho(107) = -.30, $p < .001$, and the frequency of false recognition memories for the feminine forms of high-status professions, rho(107) = -.34, $p < .001$. Finally, the frequency of false recognition memories for the feminine forms of high-status professions was directly correlated with the frequency of recognition failures for the masculine forms of high-status professions, rho(107) = .27, $p < .005$.

Taken together, these results confirm that the failures observed on this incidental memory task are not just mere errors, but reflect a systematic bias that preferentially links the masculine grammatical gender to descriptors of socially valued occupations.

Implicit Gender-Related Associations, Memory Performance, and Memory Bias

As predicted in H4 (and summarized in Table 3), the scores for the gender-status and gender-career IATs were not significantly related to any index of memory performance (correct responses, total errors, or omissions). Conversely, these D scores were correlated with the frequency of recognition failures and false recognition memories for HS stimuli, but not LS stimuli.

Regarding the gender-status IAT, it was observed that the higher the "men-high-status"/ "women-low status" associations, the higher the number of recognition failures for the feminine forms of high-status professions, rho(107) = .32, $p < .001$. Similarly, a direct correlation was observed between the gender-status IAT D scores and the higher the number of false recognition memories for the masculine forms of high-status professions in the recognition test, rho(107) = .31, p = .001. Complementarily, an inverse correlation was observed between the gender-status D scores and the forgetting the masculine forms of high-status professions, rho(107) = -.32, p < .001. Finally, the gender-status D scores were also inversely correlated to the number of false recognition memories constructed for the feminine forms of high-status professions, rho(107) = -.23, p < .05.

Similar (although slightly weaker) correlations were observed when considering the relationship between the gender-career IAT scores and the F + and F- errors for HS stimuli. More specifically, it was observed that the higher the "women-family"/ "men-professional career" associations, the higher the forgetting the feminine forms of high-status jobs, rho(107) = .25, p < .01. A direct correlation was also observed between the D scores of the gender-status IAT and the number of false memories for the masculine forms of high-status professions in the recognition test, rho(107) = .25, p < .01. Complementarily, the gender-career D scores were inversely correlated with the number of recognition failures for the masculine forms of high-status jobs, rho(107) = -.26, p < .001.

Taken together, these results suggest that the implicit gender-status associations and, to a lesser extent, the implicit gender-career associations were correlated to the memory biases observed on the incidental-encoding memory task, but not to memory capabilities/ performance.

Summary of Experiment 1 Findings

The results obtained in this experiment confirmed H1 by showing that undergraduate students harbor two sets of implicit gender-related stereotypical associations ("men-high status/ women-low status" and "men-professional career/ women-family"). The strength of these implicit stereotypes did not differ between genders.

The results of this experiment also confirmed H2 by showing that incidental encoding resulted in a low number of correct responses in the recognition test and that performance in this memory task is independent of the participants' gender. Moreover, as it had been also predicted in H2, errors in this task did not appear to be random but the result of a cognitive bias. Thus, the frequency of recognition failures was higher for the feminine forms of high-status professions (Fem-Hs) than for the masculine forms of high-status professions (Masc-Hs stimuli). Of note, the creation of false recognition memories followed the reverse and complementary pattern, so F + errors were more frequent for the masculine than for the feminine forms of high-status professions. Finally, the observed correlations between the scores of the gender-status and gender-career IATs with these different types of memory errors confirmed the predictions stated in H4 and suggest that implicit gender-related associations may be at least partially responsible for the observed memory biases.

Taken together, the results suggest that, at least when stimuli are implicitly encoded, grammatical gender may make some job-gender more or less congruent with pre-existing gender-related stereotypes and, consequently, to be processed as stereotype-consistent or stereotype-inconsistent and, therefore, end up being differentially remembered. In this regard, and because stereotype-consistent information is better recalled than stereotype-inconsistent information under conditions of reduced encoding capacity (as in our incidental memory task), but not necessarily so under normal encoding conditions, in Experiment 2 we tested whether intentional encoding is able to suppress the effects of pre-existing gender-related stereotypes on memory performance.

Experiment 2

Method

The procedure of Experiment 2 was identical to that of Experiment 1, except for two relevant aspects: First, Experiment 2 was conducted in another cohort of undergraduate students from the Universitat Jaume I (N = 96). This second sample was composed of a similar number of women (n = 49; 51.0%) and men (n = 47; 49.0%) who did not significantly differ on age (M_{men} = 18.94, SD_{men} = 0.92; M_{women} = 18.94, SD_{women} = 1.25), t(94) = -0.01, p = .991. Second, an intentional encoding memory task was used. This task was identical and involved the same stimuli and experimental parameters as the incidental encoding memory task of Experiment 1. However, in Experiment 2, participants were explicitly told that they would be participating in a memory task, and they were specifically instructed to memorize the words appearing at the center of their computer screens, hence promoting an intentional encoding of the presented stimuli. This second study was also approved by the Ethics Standards Committees of the Universitat Jaume I.

Data were analyzed using the same rationale, testing procedures, and guiding hypotheses as in Experiment 1. Moreover, because the recognition phase of the intentional memory task used in Experiment 2 involved the same number of stimuli used in the incidental-encoding memory task in Experiment 1 (i.e., 40) and the same two response options, the maximum possible number of correct responses in the intentional memory task and the number of correct responses expected by chance performance were the same as in Experiment 1 (i.e., 40 and 20, respectively).

However, consistent with H3, in this second experiment, a larger number of correct responses and a more homogeneous distribution of the different kinds of errors among the different stimuli categories (that is, attenuated memory biases) in the recognition test were expected.

Results

Implicit Gender-Career and Gender-Status Associations

H1 stated that participants would implicitly associate men with high-status attributes and women with low-status attributes. This hypothesis was tested by means of two separate one-sample Student's *t*-tests assessing whether the men's and women's mean scores in the gender-status IAT (see Table 3) were significantly larger than zero (the "no-bias" value). This prediction was confirmed for men, $t(46) = 6.58$, $p < .001$, $d = 0.96$, 95% CI [0.63, 1.41], and women, $t(48) = 5.66$, $p < .001$, $d = 0.80$, 95% CI [0.52, 1.20]. As reported in Table 4, the average D score in the gender-status IAT was slightly larger for men than women, although a Student's *t*-test for independent samples revealed that this difference was not statistically significant, $t(94) = 1.12$, $p = .26$, $d = 0.23$, 95% CI [-0.19, 0.65]. Moreover, as depicted in Fig. S1 in the online supplement, the scores in the gender-status IAT varied widely within each sex/gender category.

In a similar vein, H1 also hypothesized that men and women would implicitly associate *men* with *professional career* and *women* with *family*. Again, this hypothesis was tested employing two separate one-sample Student's *t*-tests. The obtained results confirmed that the average strength of this implicit association was significantly different from zero in men, $t(46) = 7.88$, $p < .001$, $d = 1.15$, 95% CI [0.81, 1.65], and women, $t(48) = 8.74$, $p < .001$, $d = 1.25$, 95% CI [1.00, 1.61]. In this case, the women's average IAT score was slightly larger than that observed in men (see Table 4) but, as revealed by a Student's *t*-test for independent samples, this difference did not reach statistical significance, $t(94) = -0.80$, $p = .425$, $d = -0.16$, 95% CI [-0.56, 0.26]. In addition, as illustrated in Fig. S1 in the online supplement, the magnitude of this bias largely differed between individuals of the same sex/gender category.

From these results, it can be concluded that the participants in Experiment 2 exhibited the expected stereotypical implicit gender-status associations and implicit gender-career associations. The size of these biases was similar in women and men and it was also similar to those observed in the participants recruited for Experiment 1 (see Tables 1 and 4). Thus, both samples seem to be comparable regarding their IAT scores. Therefore, if the participants in Experiment 2 exhibit a larger number of correct responses and fewer memory biases in the recognition phase of the memory task, as expected, these differences may be safely interpreted as stemming from the different encoding procedures (incidental vs. intentional) used in each experiment and not from pre-existing differences in the strength of the gender-related implicit associations exhibited by participants randomly assigned to these experimental conditions.

Table 3 Correlations Between IAT D Scores and Memory Performance and Memory Biases Indices in Experiment 1 (Incidental Encoding)

	Gender-status IAT D scores		Gender-career IAT D scores	
	rho	*p* value	rho	*p* value
Correct responses	.07	.111	.11	.252
Total errors	-.09	.413	-.14	.151
Omissions	.08	.409	.06	.541
False positive errors M-LS	.03	.789	-.10	.317
False positive errors F-LS	.04	.668	-.03	.794
False negative errors M-LS	-.12	.208	-.13	.162
False negative errors F-LS	.05	.623	.11	.250
False positive errors M-HS	**.31**	**.001**	**.25**	**.008**
False positive errors F-HS	**-.23**	**.015**	-.11	.259
False negative errors M-HS	**-.32**	**< .001**	**-.26**	**< .001**
False negative errors F-HS	**.32**	**< .001**	**.25**	**.008**

Statistically significant correlations ($p < .05$) are highlighted in bold. M-LS = masculine forms of low status professions, F-LS = feminine forms of low status professions, M-HS = masculine forms of high status professions, F-HS = feminine forms of high status professions

Memory Performance

H3 predicted that intentional encoding would result in a larger number of correct responses than that observed after incidental encoding. Confirming this hypothesis, participants of Experiment 2 averaged 29.0 correct responses, hence exhibiting a number of correct responses which is significantly larger than what could be expected by chance (20), one-sample $t(95) = 18.41$, $p < .001$, $d = 1.88$, 95% CI [1.58, 2.24], and that is also larger than the average number of correct responses observed in the incidental encoding memory task of Experiment 1 (21.9). In other words, the participants' performance in the intentional memory task equated to 72.50% of the maximum possible, which is substantially higher than the 50% expected by chance (and than the 54.75% observed after incidental encoding).

Table 4 Descriptive Statistics for the Gender-Status and Gender-Career IAT in Experiment 2

	Men		Women	
	M	(*SD*)	*M*	(*SD*)
IAT gender-status	0.28	(0.29)	0.21	(0.26)
IAT gender-career	0.28	(0.24)	0.32	(0.26)

M Mean, (*SD*) standard deviation

The second prediction of H3 was that performance in the intentional memory task would be largely independent of participant gender. This prediction was tested using a one-way MANOVA comparing the average scores of women and men for the number of correct responses, omissions, and errors. The outcome of this MANOVA, Wilks' lambda (2, 93) = 0.99, $p = .876$, as well as the results of the subsequent univariate ANOVAs (Table 5) indicated that performance in the intentional-encoding memory task was not significantly related to participant gender (see Fig. S2 in the online supplement).

As with the incidental-encoding memory task, participants could fail to recognize previously presented words (false negative errors, F-) in the intentional-encoding memory task, but they could also falsely recognize non-presented words (false positive errors, F+). Because these two kinds of errors are independent of each other, women and men could differ in the number of F- and F+ even if they did not significantly differ in the total number of errors (Table 5). To assess this possibility, a one-way MANOVA comparing the average scores of women and men on F- and F+ errors was employed. The omnibus results of this MANOVA, Wilks' lambda (2, 93) = 0.99, $p = .848$, as well as those of the subsequent one-way ANOVAs revealed that women and men incurred a similar number of F+ and F- errors in the intentional-encoding memory task.

Memory Biases

H3 stated that the number and the type of errors (recognition failures and false recognitions) would vary for the different stimuli categories, although probably less than what had been observed after the incidental encoding. This prediction was tested by means of two separate three-way ANOVAs that included the average scores of these two types of errors as dependent variables and participant gender (men vs. women), grammatical gender (Masc vs. Fem), and stimuli social value (HS v.s LS) as main factors. Two separate chi-squared goodness-of-fit tests were employed to evaluate whether recognition failures and false recognitions were evenly distributed among the four stimulus categories, respectively. The results of all these analyses are described below.

Recognition Failures (F- Errors) for Each Stimulus Category A 2 (participant gender) × 2 (grammatical gender stimuli) × 2 (social value stimuli) ANOVA comparing the average of false negative errors in the recognition test only yielded a significant effect of the interaction between the grammatical gender stimuli and the social value stimuli, $F(1,95) = 6.02$, $p < .02$, $\eta 2 = .06$. These results indicate that the specific combination of stimuli characteristics (grammatical gender and social value), but not participant gender, had a significant influence on the number of recognition failures. Thus, as revealed by subsequent posthoc comparisons (Fig. 2a), the feminine forms of high-status professions were more often forgotten than their masculine counterparts, $t(95) = 2.59$, $p_{Bonferroni} = .028$, $d = 0.38$, 95% CI [0.12, 0.66], a difference that was not reproduced for the masculine and feminine forms of low-status jobs, $t(95) = -0.79$, $p_{Bonferroni} = .868$, $d = -0.12$, 95% CI [-0.41, 0.16].

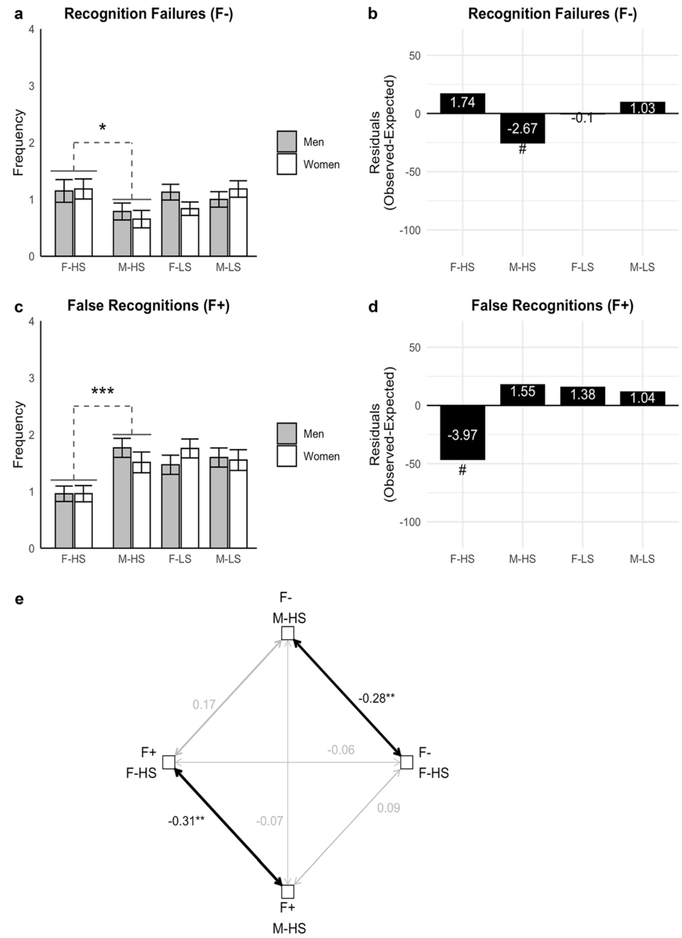
A chi-square goodness-of-fit test confirmed that recognition failures were unevenly distributed across the different stimulus categories, $\chi 2(3) = 11.22$, $p < .02$. More specifically, as Fig. 2b shows, the number of recognition failures for the masculine forms of high-status professions (was significantly lower than expected ($Z_{M\text{-}HS} = -2.67$), whereas the number of recognition failures for the feminine forms of high-status professions ($Z_{F\text{-}HS} = 1.74$) was higher than expected (although it did not surpass the critical value of ± 2; see Statistics section for details). Consequently, the probability of failing to recognize the feminine forms of high-status professions was 1.62 (95% CI [1.22, 2.16], Z = 3.32, $p < .001$) times that observed for the masculine forms of high-status professions. However, the number of F- errors for the masculine and feminine forms of low-status professions did not significantly deviate from what could be expected a priori ($Z_{F\text{-}LS} = -0.10$; $Z_{M\text{-}LS} = 1.03$) or significantly differ between each other, RR = 0.89, 95% CI [0.69, 1.16], Z = 0.82, $p = .410$. Taken together, these results confirmed that the frequency of recognition failures after intentional encoding was not evenly distributed among the four stimulus categories. This effect was similar to the one observed after incidental encoding (Experiment 1), although it appeared to be smaller in size and primarily driven by a

Table 5 Descriptive Statistics and Gender-Based Comparisons for Memory Performance Indexes in Experiment 2 (Intentional Encoding)

	Men		Women		Comparison	
	M	*(SD)*	*M*	*(SD)*	*F*(1,94)	*p*
Correct responses	28.77	(4.4)	29.22	(5.17)	0.22	.642
Total errors	9.85	(3.93)	9.63	(4.84)	0.06	.809
False positive errors (F+)	5.78	(2.56)	5.71	(2.95)	0.02	.898
False negative errors (F-)	4.06	(2.01)	3.84	(2.18)	0.28	.598
Omissions	1.38	(1.19)	1.14	(1.35)	0.85	.359

M Mean, *(SD)* Standard deviation. *F* and *p* values refer to one-way ANOVAs comparing women and men scores on memory performance indexes

Fig. 2 Memory Errors and Biases after Incidental Encoding



Note. Panels a and c depict the means ± SEM of recognition failures (false negative errors) and false recognition memories (false positive errors) displayed by women (white) and men (gray) in each stimulus category (F-HS, high-status professions in their feminine grammatical forms; M-HS, high-status professions in their masculine grammatical forms; F-LS, low-status professions in their feminine grammatical forms; M-LS, low-status professions in their masculine grammatical forms), respectively. Panels b and d depict the participants' (men and women) residual scores of false negative and false positive errors for each stimulus category, respectively. The values of the standardized residuals are provided within each bar and, in these two panels, the # symbol is used to denote standardized residuals higher than 2 or smaller than -2 and. therefore, statistically significant departures from the chance-expected number of errors. Finally, Panel e illustrates the pattern of rank order correlations (Spearman's rho) between false negative and false positive errors for the feminine (F-HS) and masculine (M-HS) forms of highly socially valued professional careers. In this panel, the arrows' width denotes the strength of the observed correlations (which exact values are provided within the figure), whereas the color denotes whether these correlations reached statistical significance (black) or not (gray). In all panels, *$p < .05$, **, $p < .01$, *** $p < .001$.

less-than-expected number of recognition failures for the masculine forms of high-status professions.

False Recognitions (F + Errors) for Each Stimulus Category A 2 (participants' gender) × 2 (grammatical gender stimuli) × 2 (social value stimuli) ANOVA revealed that the frequency of false positive errors in the recognition test was not significantly related to participant gender, $F(1, 95) = 0.01$, $p = .984$. Conversely, the number of F + errors was dependent on grammatical gender, $F(1, 95) = 9.19$, $p = .003$, $\eta 2 = .09$, and social value stimuli, $F(1, 95) = 10.17$, $p = .002$, $\eta 2 = .10$, as well as their interaction, $F(1, 95) = 9.13$, $p = .003$, η2 = .09. These results indicate that the specific combination of stimuli characteristics (grammatical gender and social value), but not participant gender, had a statistically significant influence on the number of F + errors. Thus, as Fig. 3c shows, participants were less likely to erroneously recognize the feminine compared to the masculine forms of non-previously presented high-status professions, $t(95) = -4.11$, $p_{Bonferroni} < .001$, $d = 0.62$, 95% CI [-0.9, -0.35]. In contrast, the frequency of F + errors did not differ in a statistically significant manner between the masculine and feminine forms of low-status jobs, $t(95) = -0.27$, $p_{Bonferroni} = 1$, $d = 0.04$, 95% CI [-0.25, 0.34].

The same conclusions were reached after performing a chi-squared goodness-of- fit test. This test revealed an uneven distribution of false recognitions across the four stimulus categories, $\chi 2(3) = 21.14$, $p < .001$ (see Fig. 3d), and the inspection of the standardized residuals revealed that the frequency of this kind of error for the feminine form of high-status professions was significantly lower than expected ($Z_{F\text{-}HS} = -3.97$). In contrast, the number of F + errors for the rest of the stimulus categories was not significantly different from what was expected ($Z_{M\text{-}HS} = 1.55$, $Z_{F\text{-}LS} = 1.38$, and $Z_{M\text{-}LS} = 1.04$, respectively). Thus, the risk of erroneously recognizing feminine forms of high-status professions that had not been previously presented was about 60% of that observed for their masculine counterparts, RR = 0.59, 95% CI [0.46, 0.75], $Z = 4.34$, $p < .001$, whereas the risk of F + errors on the feminine and masculine forms of low status professions was not significantly different (RR = 0.98, 95% CI [0.79, -1.20], $Z = 0.23$, $p = .814$). Taken together, these results indicate that false recognition memories were not random but biased for some stimuli categories. This effect was similar to the one observed after incidental encoding (Experiment 1), although it appeared to be smaller in size and solely driven by the tendency of participants to create fewer than expected false recognition memories for feminine forms of high-status professions.

Relationship Between F- and F + Errors on HS Stimuli From the results described above, it can be concluded that intentional encoding results in a reduced number of errors in the recognition test, but also that these errors are not entirely random. This suggests that, after intentional encoding, there is a limited yet still significant influence of grammatical gender on the frequency and distribution of recognition failures (F-) and false recognition memories (F+) for socially valued occupations (HS stimuli).

In agreement with this interim conclusion, we observed that intentional encoding partially disrupted the pattern of correlations between F-/ F + errors for stimuli referring to masculine and feminine forms of high-status professions observed on the incidental-encoding memory task (Fig. 1e). In fact, as Fig. 3e illustrates, after intentional encoding, the frequency of F- errors on the feminine forms of high status professions and the frequency of F- errors on the masculine forms of high status professions remained inversely correlated, rho(94) = -.28, $p < .005$. A similar inverse correlation was observed between the number of F + errors on the feminine forms of high status professions and the number of F + errors on the masculine forms of high status professions, rho(94) = -.31, $p < .002$. These correlations are weaker than those observed in Experiment 1, an observation that is in agreement with the smaller effects also observed in the relative distribution of F + and F- errors on F-HS and M-HS stimuli after intentional encoding. Taken together, these results suggest that intentional encoding largely erodes, but does not totally suppress, the influence of gender-status stereotypes on memory retrieval.

Implicit Associations, Memory Performance, and Memory Bias

H4 stated that intentional encoding would weaken or even totally suppress the relationship between the gender-status and gender-career scores and the frequency of recognition failures (F-) and false recognition memories (F+) for HS stimuli observed after incidental encoding (see Fig. 2). On the other hand, as in Experiment 1, no correlations between the scores of these IATs and memory performance indexes should be expected.

In agreement with these predictions, the D scores on the gender-status and gender-career IATs were not significantly related to memory performance (Table 6). Moreover, most of the correlations between these IAT scores and the frequency of recognition failures and false recognition memories for HS stimuli observed in Experiment 1 (see Table 3) faded away. However, it is worth noting that the D scores on the IAT gender-status were inversely/ directly correlated with the relative frequency of false recognition memories for the feminine forms of high status professions, rho(94) = -0.2, $p < .05$, and the masculine forms of high status professions, rho(94) = 0.26, $p < .05$, respectively. In addition, the D scores on the IAT gender-career were statistically and directly correlated with the relative frequency of recognition failures for the feminine forms of high status professions, rho(94) = 0.24, $p < .05$.

Therefore, these findings suggest that intentional memory encoding largely reduced, but did not eliminate, the effects of implicit gender-related associations on the differential occurrence of F- and F + errors for the feminine and masculine forms of HS stimuli.

Summary of Experiment 2 Findings

The results obtained in experiment 2 re-confirmed H1 by showing again that, regardless their gender, undergraduate students harbor implicit associations that preferentially link men to high status and professional career and women to low status and family.

The results of this experiment also revealed that intentional encoding resulted in a larger number of correct response s in the recognition test than incidental encoding, then confirming the first proposition of H3. Moreover, as it had been also hypothesized in H3, intentional encoding attenuated -although did not totally suppress- memory biases triggered by the grammatical gender of stimuli referring to high status professions. In agreement with this finding, the correlations between these errors and the scores of the gender-status and gender-career IATs were attenuated and several of them did not achieve statistical significance, hence confirming the predictions of H4.

General Discussion

The results of this research suggest that people implicitly associate gender and social status, and that this association is incorporated into gender schemata, leading to biased information processing. This bias became more prominent on our incidental memory task, where selective memory failures and false recognitions formed a coherent pattern of F + and F- errors that was consistently correlated with the strength of the implicit association that selectively links men to high status professions.

More specifically, the results obtained on the self-constructed IAT gender-status show that, regardless of gender, participants implicitly associated "men" with "high status" and/or "women" with "low status" attributes. This association has a similar size and it is partially correlated with another well-established gender-related implicit association linking *men* to *professional career* and *women* to *family* (the gender-career IAT; Nosek et al., 2002). To our knowledge, this gender-status association had not been directly and explicitly assessed but our results converge with previous studies using the IAT that have found an association between gender categories and professional occupations that differ in authority, economic retribution, and/or social prestige (e.g., Levinson & Young, 2010; Matheus & Quinn, 2017; White & White, 2006).

For instance, Levinson and Young (2010) employed a self-developed "judge/gender" IAT to show that people associate high-status positions in legal professions (e.g., judge) with *men*, and lower status positions (e.g., paralegal) with *women*. Moreover, as in our study, Levinson and Young (2010) observed that their participants also held a "men-professional career/ women-family" association, and that these constructs were partially related, concluding that these associations tap two related but different gender-related stereotypes, consistent with our studies. In addition, our results also converge with findings that show the best predictor of salary and prestige of a particular job is the degree to which it is perceived as requiring traits stereotypically associated with men (Glick, 1991; Glick et al., 1995). Together, these studies support the notion that there is a gender-status association and that it fosters the perception of prestigious and socially valued professions as masculine. Of note, as in our study, none of these previous investigations reported an effect of the participants' sex/gender (and, in some cases, they also ruled out age and educational level effects), which suggests that the gender-status association is probably shared by many different socio-demographic groups.

However, the main finding of our study was that an implicit gender-status association is effectively incorporated into gender schemata, biasing information processing and affecting some cognitive processes, such as memory. Thus, in our incidental-encoding memory task, participants were less likely to remember the feminine forms of high-status occupations than any other kind of stimuli, whereas the exact opposite was observed for the masculine forms of the same professions, which were forgotten least often. Moreover, participants were more/less prone to falsely recognize the masculine and the feminine forms of high-status professions, respectively, than any other stimulus category. The frequencies of these four kinds of errors were not only significantly different from what would be expected by chance, but they were also coherently correlated with each other. In other words, the frequency of F + errors for the masculine forms of high status professions was directly correlated with the frequency of F- errors for the feminine forms of high status professions, but it was also inversely correlated with the number of F- errors on for the masculine forms of high status professions and F + errors on the feminine forms of high status professions, which, in turn, were directly/inversely correlated with the number of F- errors on the masculine forms of high status professions and the feminine forms of high status professions, respectively.

Therefore, the memory errors observed on this task were not random, but rather they formed an internally consistent pattern. Furthermore, these errors were significantly correlated and in tune with the strength of the implicit gender-status association measured by the IAT. Thus, high IAT gender-status scores (which indicate strong men-high-status/ women-low status associations) were associated with a larger number of false memories for the masculine forms of high-status professions and forgetting of the feminine forms of high status professions, as well as with a lower number of false memories for the feminine forms of high status professions and forgetting of the masculine forms of high status professions.

Taken together, the results obtained on the incidental-encoding memory task suggest that grammatical gender is able to inadvertently convey the content of the implicit association which stereotypically associates high status attributes with men and/ or low status with women, making the masculine forms of high-status professions more likely to be processed as stereotype-consistent and the feminine forms of high status professions as stereotype-inconsistent information. Consequently, participants created false memories about the masculine forms of high-status professions and suppressed true memories of the feminine forms of high status professions, and so their memory recollection became consistent with their pre-existing gender schemata.

These observations suggest that information about a person's sex/gender through grammatical gender markers may make some job-sex combinations more acceptable (that is, more congruent with pre-existing gender-status stereotypes) and, consequently, processed as stereotype-consistent or stereotype-inconsistent and differentially remembered. In this way, our findings and interpretation concur with those of previous studies showing that the gender cues present in words (Bem & Bem, 1973; Stahlberg et al., 2007; Vervecken et al., 2013) or in graphical images (Liben et al., 2001) describing occupations influence the perception and valuation of the described occupations. Furthermore, the present study extends these previous findings by showing that, at least in some cases, the information conveyed by gender cues specifically refers to the perceived relationship between gender and social status.

Table 6 Correlations Between IAT D Scores and Memory Performance and Memory biases Indices in Experiment 2 (Intentional Encoding)

	Gender-status IAT D scores		Gender-career IAT D scores	
	rho	*p* value	rho	*p* value
Correct responses	-.09	.401	.14	.187
Total errors	.06	.581	-.12	.230
Omissions	-.04	.422	-.03	.807
False positive errors M-LS	-.05	.657	-.06	.586
False positive errors F-LS	-.01	.954	-.08	.422
False negative errors M-LS	0.0	.996	-.12	.233
False negative errors F-LS	.09	.388	-.10	337
False positive errors M-HS	**.26**	**.011**	.11	.230
False positive errors F-HS	**-.20**	**.049**	-.01	351
False negative errors M-HS	.08	.410	-.14	.173
False negative errors F-HS	-.09	.361	**.24**	**.016**

Statistically significant correlations ($p < .05$) are highlighted in bold. M-LS = masculine forms of low status professions, F-LS = feminine forms of low status professions, M-HS = masculine forms of high status professions, F-HS = feminine forms of high status professions

On the other hand, our results and conclusions are also coincident with those of studies showing that children (e.g.,Carter & Levy, 1988; Frawley, 2008; Levy, 1989; Liben & Signorella, 1980, 1993; Signorella & Liben, 1984; Stangor & Ruble, 1989) and adults (e.g., Bem, 1981; Cherney, 2005; Herrmann et al., 1992; Renn & Calvert, 1993; Shapiro, 2009) have better memory for gender schemata-consistent materials than for gender schemata-inconsistent information. Similarly, our results also align with evidence specifically showing that stereotypes can promote false memories that are congruent with pre-existing gender-related beliefs (e.g., Kleider et al., 2008; Lenton et al., 2001; MacRae et al., 2010; Stangor, 1988), and that stereotype-congruent memories are more difficult to suppress than non-stereotypical memories (MacRae et al., 1993). Most of these studies have also shown that memory distortions induced by gender-stereotypes are more frequent in gender-typed individuals, an observation that aligns with the correlations between the IAT scores and the pattern of memory errors identified in our memory task.

Of note, most effects produced by gender stereotypes on the incidental-encoding memory task (Experiment 1) decayed when an intentional-encoding memory task was used (Experiment 2). In fact, when participants were explicitly instructed to memorize as many stimuli as possible, memory errors became more evenly distributed across all the stimulus categories (i.e., the frequency of F + errors on the masculine forms of high status professions was not different from that observed for the masculine forms of low status professions), and unlike what was observed on the incidental-encoding memory task, only some of them formed a coherent pattern (i.e., the correlations between most types of memory errors were weak and they did not reach statistical significance). In a similar vein, most but not all correlations between IATs scores and memory errors for the feminine and masculine forms of high-status professions stimuli decayed, reinforcing the notion that some of them were probably true errors whereas others probably were the result of an implicit stereotypes-related bias. All these findings support the predictions of gender schemata theory on the effects of stereotypes on memory. In this regard, it has been shown that stereotype-consistent information is better recalled than stereotype-inconsistent information under conditions of reduced encoding capacity (as in our incidental memory task), but not under normal encoding conditions (e.g., Bodenhausen & Lichtenstein, 1987; Cherney, 2005; MacRae et al., 1993; Stangor & Duan, 1991; Stangor & McMillan, 1992). Similarly, it has been shown that divided or misguided attention during encoding can lead to reduced memory performance (e.g., McLaughlin, 1965; Naveh-Benjamin et al., 2014), and stereotype-consistent false memories (MacRae et al., 2010).

Limitations and Future Directions

Our findings and conclusions must be considered in light of the limitations of our study. First, although the size of the samples used in the present study cannot be considered small, it might have been insufficient to detect some small-sized effects. This might be especially relevant for the results of Experiment 2, in which the number of errors (and, consequently, the chances to identify biases) are reduced. In a similar vein, it was this limited sample size that prevented us from conducting a formal comparison between the incidental and the intentional experimental conditions, an issue that should be directly addressed by future studies.

Second, the standard IAT used in the present study only provides limited information about the measured associations. Thus, because it uses complementary pairs of concepts and attributes, this standard procedure precludes ascertaining whether the contents of the gender-status implicit association identified in the present study is actually relating men to high-status, women to low-status, or both. However, future studies using the single-category IAT (Sc-IAT; Karpinski & Steinman, 2006) could be used to address this question. In addition, the autobiographical IAT (aIAT; Agosta & Sartori, 2013) could be used to distinguish between false and true autobiographical remembrances for feminine and masculine forms of high/low-status professions (see related applications in Marini et al., 2012; Sartori et al., 2008).

Third, our study shows that the implicit gender-status association (and, to a lesser extent, the gender-career association) results in a different processing/ remembering of the masculine and feminine forms of nouns referring to socially valued professions. Based on theory and previous findings, it might be hypothesized that this differential processing should affect the perception and valuation of these jobs and, thereby, academic and professional choices (Bem & Bem, 1973; Born & Taris, 2010; Gaucher et al., 2011; Liben et al., 2002; Stout & Dasgupta, 2011; Vervecken et al., 2013). However, whether and how these predicted consequences do take place requires further investigation.

Practice Implications

The results of the present study show implicit gender-status associations are incorporated into gender schemata and result in the biased processing of occupational titles (i.e., decreased forgetting and increased false memories for the masculine forms of high-status professions and the reverse pattern for their feminine counterparts). While these findings primarily contribute to enhancing our knowledge about how the contents of gender schemata impact the processing of information, it seems reasonable to suggest that this implicit association and its effects on cognition may lead to the justification, normalization, and perpetuation of gender-based asymmetries in the workplace.

In this regard, social role theory proposes that gender stereotypes are not solely a consequence of gender-based divisions of labor but also one of the major forces that maintain them (Eagly et al., 2000). Moreover, research in this area has revealed that implicit stereotypical associations are especially likely to satisfy the individuals' generalized psychological tendency to justify/ accept the status quo and, therefore, to reinforce the realities from which these stereotypes initially arose (Jost & Hamilton, 2005; Jost et al., 2004). Thus, it can be proposed that, in the same way that it promoted a selective forgiveness of stereotype-inconsistent stimuli, the implicit association between men with high status attributes could make individuals blind to the fact that they are being privileged or discriminated because of their gender but also contribute to the perception/ internalization of these gender-based asymmetries as natural, appropriate, or justified, and/or promote a non-conscious readjustment of their expectations to fit the status quo. Accordingly, this association seems potentially relevant to explain gender differences in pay entitlement (Blanton et al., 2001; Desmarais & Curtis, 1997; O'Brien, 2012; Pelham & Hetts, 2001) and the persistent interest of girls in female-typed jobs despite knowing they offer less social prestige and lower salaries (Liben et al., 2001; Weisgram et al., 2010). Similarly, this association may be useful to understand why women occupying high status professions can be perceived as "illegitimate intruders" by others and/ or as "impostors" by themselves (Eagly & Karau, 2002; Young, 2011), as well as explain why gender equality-policies are so often resisted and backlashed (Liquat et al., 2023).

Conclusion

In conclusion, our study shows that a contemporary sample of undergraduate students endorse gender-related stereotypes about occupations. More importantly, our study also shows that these stereotypes are integrated into gender schemas that can be activated by grammatical gender markers and that produce a consistent bias on memory and, probably, other cognitive functions. Although these memory biases were more evident under incidental-encoding conditions, they also seemed to unwittingly leak into the voluntary processing of new incoming information, hence interfering with participants' conscious attempt to memorize a list of professional occupations. This observation speaks to the pervasiveness and persistence of the effects of gender stereotypes on cognitive processes even in situations in which gender is not really relevant. On the other hand, these effects were only observed for highly valued professions, thus highlighting that the contents of these professional occupation gender stereotypes are related to a social status dimension that combines prestige, authority, and economic retribution that may contribute to explaining the persistence of several forms of gender inequality in the workplace. Overall, our results indicate that, despite the progress made toward gender equality in recent decades, gender stereotypes may still represent a psychological barrier to full gender equality.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11199-023-01355-z.

Acknowledgements This research was supported by a grant (PID2019-106793RB-I00/AEI/10.13039/501100011033) provided by Ministerio de Ciencia e Innovación to CF and CS-S and a grant (UJI B2020-02) awarded to CF and CS-S. S.F. was supported by an FPI grant from UJI (PREDOC/2020/22).

Authors' Contribution C.F. conceived the study. C.F. and C.S–S. designed and conceptualized the study. All authors participated in the experimental sessions. A.S-T. and S.F preprocessed the data with which C.S.-S. conducted the statistical analyses. C.S.-S. and C.F. wrote the manuscript. All authors contributed to manuscript revision and read, edited, and approved the submitted version.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. This research was supported by a grant (PID2019-106793RB-I00/AEI/10.13039/501100011033) provided by Ministerio de Ciencia e Innovación to CF and CS-S and a grant (UJI B2020-02) awarded to CF and CS-S. S.F. was supported by an FPI grant from UJI (PREDOC/2020/22).

Data Availability The data that support the findings of this study are available from the corresponding author upon request and after receiving the authorization of Universitat Jaume I.

Compliance with Ethical Standards

Ethical Approval The study was approved by the Ethics Standards Committees of the Universitat Jaume I. All participants signed informed consents.

Competing Interests The authors declare no competing interests.

References

Agosta, S., & Sartori, G. (2013). The autobiographical IAT: A review. *Frontiers in Psychology, 4*, 519. https://doi.org/10.3389/fpsyg.2013.00519

Agresti, A. (2007). An introduction to categorical data analysis. *Wiley*. https://doi.org/10.1002/0470114754

Alksnis, C., Desmarais, S., & Curtis, J. (2008). Workforce segregation and the gender wage gap: Is "women's" work valued as highly as "men's"? *Journal of Applied Social Psychology, 38*(6), 1416–1441. https://doi.org/10.1111/j.1559-1816.2008.00354.x

Atli, A. (2017). High school students' gender role perceptions regarding various professions. *International Journal of Progressive Education, 13*(3), 6–15.

Beggs, J. M., & Doolittle, D. C. (1993). Perceptions now and then of occupational sex typing: A replication of shinar's 1975 study. *Journal of Applied Social Psychology, 23*(17), 1435–1453. https://doi.org/10.1111/j.1559-1816.1993.tb01042.x

Bem, S. L. (1981). Gender schema theory: A cognitive account of sex typing. *Psychological Review, 88*(4), 354–364. https://doi.org/10.1037/0033-295X.88.4.354

Bem, S. L. (1983). Gender schema theory and its implications for child development: Raising gender-aschematic children in a gender-schematic society. *Signs, 8*, 598–616. https://doi.org/10.1086/493998

Bem, S. L. (1993). *The lenses of gender: Transforming the debate on sexual inequality*. Yale University Press.

Bem, S. L., & Bem, D. J. (1973). Does sex-biased job advertising "aid and abet" sex discrimination? *Journal of Applied Social Psychology, 3*(1), 6–18. https://doi.org/10.1111/j.1559-1816.1973.tb01290.x

Blair-Loy, M. (2009). *Competing devotions: Career and family among women executives*. Harvard University Press.

Blanton, H., George, G., & Crocker, J. (2001). Contexts of system justification and system evaluation: Exploring the social comparison strategies of the (not yet) contented female worker. *Group Processes & Intergroup Relations, 4*(2), 126–137. https://doi.org/10.1177/1368430201004002004

Bodenhausen, G. V., & Lichtenstein, M. (1987). Social stereotypes and information-processing strategies: The impact of task complexity. *Journal of Personality and Social Psychology, 52*(5), 871. https://doi.org/10.1037/0022-3514.52.5.871

Born, M. P., & Taris, T. W. (2010). The impact of the wording of employment advertisements on students' inclination to apply for a job. *The Journal of Social Psychology, 150*(5), 485–502. https://doi.org/10.1080/00224540903365422

Callahan-Levy, C. M., & Messé, L. A. (1979). Sex differences in the allocation of pay. *Journal of Personality and Social Psychology, 37*(3), 433. https://doi.org/10.1037/0022-3514.37.3.433

Carreiras, M., Garnham, A., Oakhill, J., & Cain, K. (1996). The use of stereotypical gender information in constructing a mental model: Evidence from english and spanish. *The Quarterly Journal of Experimental Psychology A, 49*(3), 639–663. https://doi.org/10.1080/027249896392531

Carter, D. B., & Levy, G. D. (1988). Cognitive aspects of early sex-role development: The influence of gender schemas on preschoolers' memories and preferences for sex-typed toys and activities. *Child Development, 59*(3), 782. https://doi.org/10.2307/1130576

Charles, M., & Grusky, D. B. (2005). *Occupational ghettos: The worldwide segregation of women and men* (Vol. 200). Stanford University Press.

Cherney, I. D. (2005). Children's and adults' recall of sex-stereotyped toy pictures: Effects of presentation and memory task. *Infant and Child Development, 14*(1), 11–27. https://doi.org/10.1002/icd.372
Ciani, K. D., Summers, J. J., & Easter, M. A. (2008). Gender differences in academic entitlement among college students. *The Journal of Genetic Psychology, 169*(4), 332–344. https://doi.org/10.3200/GNTP.169.4.332-344
Clarke, H. M. (2020). Gender stereotypes and gender-typed work. *Handbook of Labor, Human Resources and Population Economics*, 1–23. https://doi.org/10.1007/978-3-319-57365-6_21-1
Cohen, P., Cohen, J., Aiken, L. S., & West, S. G. (1999). The problem of units and the circumstance for POMP. *Multivariate Behavioral Research, 34*(3), 315–346. https://doi.org/10.1207/S15327906MBR3403_2
Correll, S. J., Benard, S., & Paik, I. (2014). Is there a motherhood penalty?. *Social Stratification: Class, Gender, and Race in Sociological Perspective*. https://doi.org/10.1086/511799
Cotter, D. A., Hermsen, J. M., Ovadia, S., & Vanneman, R. (2001). The glass ceiling effect. *Social Forces, 80*(2), 655–681. https://doi.org/10.1353/sof.2001.0091
Crosby, F. (1982). *Relative deprivation and working women*. Oxford University Press.
Dasgupta, N., & Asgari, S. (2004). Seeing is believing: Exposure to counterstereotypic women leaders and its effect on the malleability of automatic gender stereotyping. *Journal of Experimental Social Psychology, 40*(5), 642–658. https://doi.org/10.1016/j.jesp.2004.02.003
Desmarais, S., & Curtis, J. (1997). Gender differences in pay histories and views on pay entitlement among university students. *Sex Roles, 37*(9/10), 623. https://doi.org/10.1007/BF02936332
Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review, 109*(3), 573. https://doi.org/10.1037/0033-295X.109.3.573
Eagly, A. H., & Mladinic, A. (1994). Are people prejudiced against women? Some answers from research on attitudes, gender stereotypes, and judgments of competence. *European Review of Social Psychology, 5*(1), 1–35. https://doi.org/10.1080/14792779543000002
Eagly, A. H., Wood, W., & Diekman, A. B. (2000). Social role theory of sex differences and similarities: A current appraisal. In: *The developmental social psychology of gender*. Edited By Thomas Eckes, Hanns M. Trautner Psychology Press. ISBN: 9781410605245.
Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191. https://doi.org/10.3758/BF03193146
Frawley, T. J. (2008). Gender schema and prejudicial recall: How children misremember, fabricate, and distort gendered picture book information. *Journal of Research in Childhood Education, 22*(3), 291–303. https://doi.org/10.1080/02568540809594628
García-Ael, C., Cuadrado, I., & Molero, F. (2018). The effects of occupational status and sex-typed jobs on the evaluation of men and women. *Frontiers in Psychology, 9*, 1170. https://doi.org/10.3389/fpsyg.2018.01170
Gaucher, D., Friesen, J., & Kay, A. C. (2011). Evidence that gendered wording in job advertisements exists and sustains gender inequality. *Journal of Personality and Social Psychology, 101*(1), 109. https://doi.org/10.1037/a0022530
Glick, P. (1991). Trait-based and sex-based discrimination in occupational prestige, occupational salary, and hiring. *Sex Roles, 25*(5), 351–378. https://doi.org/10.1007/BF00289761
Glick, P., Wilk, K., & Perreault, M. (1995). Images of occupations: Components of gender and status in occupational stereotypes. *Sex Roles, 32*(9), 565–582. https://doi.org/10.1007/BF01544212
Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review, 109*(1), 3. https://doi.org/10.1037/0033-295X.109.1.3
Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74*(6), 1464–1480. https://doi.org/10.1037/0022-3514.74.6.1464
Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology, 85*(2), 197–216. https://doi.org/10.1037/0022-3514.85.2.197
Guo, Y., Guo, Y., & He, X. (2015). A stereotypic explanatory bias study of college students' occupational gender stereotype. *Canadian Social Science, 11*(3), 141–145. https://doi.org/10.3968/6468
Halper, L. R., Cowgill, C. M., & Rios, K. (2019). Gender bias in caregiving professions: The role of perceived warmth. *Journal of Applied Social Psychology, 49*(9), 549–562. https://doi.org/10.1111/jasp.12615
Hegewisch, A., Liepmann, H., Hayes, J., & Hartmann, H. (2010). Separate and not equal? Gender segregation in the labor market and the gender wage gap. *IWPR Briefing Paper*, *377*, 1–16. https://www.jstor.org/stable/pdf/resrep36063.pdf
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
Herrmann, D. J., Crawford, M., & Holdsworth, M. (1992). Gender-linked differences in everyday memory performance. *British Journal of Psychology, 83*(2), 221–231. https://doi.org/10.1023/A:1022082023501
Jost, J. T., Banaji, M. R., & Nosek, B. A. (2004). A decade of system justification theory: Accumulated evidence of conscious and unconscious bolstering of the status quo. *Political Psychology, 25*(6), 881–919. https://doi.org/10.1111/j.1467-9221.2004.00402.x
Jost, J. T., & Hamilton, D. L. (2005). Stereotypes in our culture. *On the nature of prejudice: Fifty years after Allport*, 208–224. https://doi.org/10.1002/9780470773963.ch13
Karpinski, A., & Steinman, R. B. (2006). The single category implicit association test as a measure of implicit social cognition. *Journal of Personality and Social Psychology, 91*(1), 16. https://doi.org/10.1037/0022-3514.91.1.16
Kleider, H., Goldinger, S., & Knuycky, L. (2008). Stereotypes influence false memories for imagined events. *Memory, 16*(2), 97–114. https://doi.org/10.1080/09658210801895948
Koenig, A. M., & Eagly, A. H. (2014). Evidence for the social role theory of stereotype content: Observations of groups' roles shape stereotypes. *Journal of Personality and Social Psychology, 107*(3), 371. https://doi.org/10.1037/a0037215
Lenton, A. P., Blair, I. V., & Hastie, R. (2001). Illusions of gender: Stereotypes evoke false memories. *Journal of Experimental Social Psychology, 37*(1), 3–14. https://doi.org/10.1006/jesp.2000.1426
Leung, E. (2020). Gender schemas. In Zeigler-Hill, V & Shackelford, T. K (Eds.), *Encyclopedia of personality and individual differences.* Springer. https://doi.org/10.1007/978-3-319-24612-3_667
Levanon, A., & Grusky, D. B. (2016). The persistence of extreme gender segregation in the twenty-first century. *American Journal of Sociology, 122*(2), 573–619. https://doi.org/10.1086/688628
Levinson, J. D., & Young, D. (2010). Implicit gender bias in the legal profession: An empirical study. *Duke J. Gender L. & Pol'y, 18*, 1.
Levy, G. D. (1989). Developmental and individual differences in preschoolers' recognition memories: The influences of gender schematization and verbal labeling of information. *Sex Roles, 21*(5), 305–324. https://doi.org/10.1007/BF00289594

Liaquat, U., Jost, J. T., & Balcetis, E. (2023). System justification motivation as a source of backlash against equality-promoting policies—and what to do about it. *Social Issues and Policy Review*. https://doi.org/10.1111/sipr.12093

Liben, L. S., Bigler, R. S., & Krogh, H. R. (2001). Pink and blue collar jobs: Children's judgments of job status and job aspirations in relation to sex of worker. *Journal of Experimental Child Psychology, 79*(4), 346–363. https://doi.org/10.1006/jecp.2000.2611

Liben, L. S., Bigler, R. S., & Krogh, H. R. (2002). Language at work: Children's gendered interpretations of occupational titles. *Child Development, 73*(3), 810–828. https://doi.org/10.1111/1467-8624.00440

Liben, L. S., & Signorella, M. L. (1980). Gender-related schemata and constructive memory in children. *Child Development, 51*(1), 11. https://doi.org/10.2307/1129584

Liben, L. S., & Signorella, M. L. (1993). Gender-schematic processing in children: The role of initial interpretations of stimuli. *Developmental Psychology, 29*(1), 141. https://doi.org/10.1037/0012-1649.29.1.141

Lips, H., & Lawson, K. (2009). Work values, gender, and expectations about work commitment and pay: Laying the groundwork for the "motherhood penalty"? *Sex Roles, 61*(9), 667–676. https://doi.org/10.1007/s11199-009-9670-0

MacRae, C. N., Hewstone, M., & Griffiths, R. J. (1993). Processing load and memory for stereotype-based information. *European Journal of Social Psychology, 23*(1), 77–87. https://doi.org/10.1002/ejsp.2420230107

MacRae, C. N., Schloerscheidt, A. M., Bodenhausen, G. V., & Milne, A. B. (2010). Creating memory illusions: Expectancy-based processing and the generation of false memories. *Memory, 10*(1), 63–80. https://doi.org/10.1080/09658210143000254

Major, B. (1994). From social inequality to personal entitlement: The role of social comparisons, legitimacy appraisals, and group membership. In *Advances in experimental social psychology* (Vol. 26, pp. 293–355). Academic Press. https://doi.org/10.1016/S0065-2601(08)60156-2

Major, B., & Forcey, B. (1985). Social comparisons and pay evaluations: Preferences for same-sex and same-job wage comparisons. *Journal of Experimental Social Psychology, 21*(4), 393–405. https://doi.org/10.1016/0022-1031(85)90038-1

Major, B., McFarlin, D. B., & Gagnon, D. (1984). Overworked and underpaid: On the nature of gender differences in personal entitlement. *Journal of Personality and Social Psychology, 47*(6), 1399. https://doi.org/10.1037/0022-3514.47.6.1399

Marini, M., Agosta, S., Mazzoni, G., Barba, G. D., & Sartori, G. (2012). True and false DRM memories: Differences detected with an implicit task. *Frontiers in Psychology, 3*, 310. https://doi.org/10.3389/fpsyg.2012.00310

Marini, M., & Banaji, M. R. (2022). An implicit gender sex-science association in the general population and STEM faculty. *The Journal of General Psychology, 149*(3), 299–326. https://doi.org/10.1080/00221309.2020.1853502

Martin, C. L., Ruble, D. N., & Szkrybalo, J. (2002). Cognitive theories of early gender development. *Psychological Bulletin, 128*(6), 903. https://doi.org/10.1037/0033-2909.128.6.903

Martin, A. E. (2022). Gender relativism: How context shapes what is seen as male and female. *Journal of Experimental Psychology: General*. Advance online publication. https://doi.org/10.1037/xge0001264

Matheus, C. C., & Quinn, E. (2017). Gender based occupational stereotypes: New behaviors, old attitudes. *2017 IEEE Women in Engineering (WIE) Forum USA East*, 1–6. https://doi.org/10.1109/WIE.2017.8285610

McLaughlin, B. (1965). "Intentional" and "incidental" learning in human subjects: The role of instructions to learn and motivation. *Psychological Bulletin, 63*(5), 359–376. https://doi.org/10.1037/H0021759

Naveh-Benjamin, M., Guez, J., Hara, Y., Brubaker, M. S., & Lowenschuss-Erlich, I. (2014). The effects of divided attention on encoding processes under incidental and intentional learning instructions: Underlying mechanisms? *Quarterly Journal of Experimental Psychology, 67*(9), 1682–1696. https://doi.org/10.1080/17470218.2013.867517

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Harvesting implicit group attitudes and beliefs from a demonstration web site. *Group Dynamics: Theory, Research, and Practice, 6*(1), 101. https://doi.org/10.1037/1089-2699.6.1.101

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2007). The Implicit Association Test at age 7: A methodological and conceptual review. In J. A. Bargh (Ed.), *Social psychology and the unconscious: The automaticity of higher mental processes* (pp. 265–292). Psychology Press.

Nosek, B. A., Smyth, F. L., Sriram, N., Lindner, N. M., Devos, T., Ayala, A., Bar-Anan, Y., Bergh, R., Cai, H., Gonsalkorale, K., Kesebir, S., Maliszewski, N., Neto, F., Olli, E., Park, J., Schnabel, K., Shiomura, K., Tulbure, B. T., Wiers, R. W., & Greenwald, A. G. (2009). National differences in gender–science stereotypes predict national sex differences in science and math achievement. *Proceedings of the National Academy of Sciences, 106*(26), 10593–10597. https://doi.org/10.1073/pnas.0809921106

Oakhill, J., Garnham, A., & Reynolds, D. (2005). Immediate activation of stereotypical gender information. *Memory & Cognition, 33*(6), 972–983. https://doi.org/10.3758/BF03193206

O'Brien, L. T., Major, B. N., & Gilbert, P. N. (2012). Gender differences in entitlement: The role of system-justifying beliefs. *Basic and Applied Social Psychology, 34*(2), 136–145. https://doi.org/10.1080/01973533.2012.655630

Pelham, B. W., & Hetts, J. J. (2001). Underworked and overpaid: Elevated entitlement in men's self-pay. *Journal of Experimental Social Psychology, 37*(2), 93–103. https://doi.org/10.1006/jesp.2000.1429

Pyykkönen, P., Hyönä, J., & Van Gompel, R. P. G. (2010). Activating gender stereotypes during online spoken language processing: Evidence from visual world eye tracking. *Experimental Psychology, 57*(2), 126–133. https://doi.org/10.1027/1618-3169/A000016

Renn, J. A., & Calvert, S. L. (1993). The relation between gender schemas and adults' recall of stereotyped and counterstereotyped televised information. *Sex Roles, 28*(7), 449–459. https://doi.org/10.1007/BF00289607

Ridgeway, C. L. (2011). Framed by gender: How gender inequality persists in the modern world. *Oxford University Press*. https://doi.org/10.1177/0891243212439344

Ridgeway, C. L., & Erickson, K. G. (2000). Creating and spreading status beliefs. *American Journal of Sociology, 106*(3), 579–615. https://doi.org/10.1086/318966

Salanauskaite, D. L. (March, 2017). *Gender segregation in education, training and the labour market: Emerging findings from the Beijing Platform for Action report*. STEM Gender Equality Congress. https://pdfs.semanticscholar.org/664d/30650c495c3248ba69d800dd30d8b2db1f7c.pdf

Salles, A., Awad, M., Goldin, L., Krus, K., Lee, J. V., Schwabe, M. T., & Lai, C. K. (2019). Estimating implicit and explicit gender bias among health care professionals and surgeons. *JAMA Network Open, 2*(7), Article e196545–e196545. https://doi.org/10.1001/jamanetworkopen.2019.6545

Sartori, G., Agosta, S., Zogmaister, C., Ferrara, S. D., & Castiello, U. (2008). How to accurately detect autobiographical events. *Psychological Science, 19*(8), 772–780. https://doi.org/10.1111/j.1467-9280.2008.02156.x

Schnabel, K., Asendorpf, J. B., & Greenwald, A. G. (2008). Understanding and using the implicit association test: V. measuring semantic aspects of trait self-concepts. *European Journal of Personality: Published for the European Association of Personality Psychology, 22*(8), 695–706. https://doi.org/10.1002/per.697

Sellers, N., Satcher, J., & Comas, R. (1999). Children's occupational aspirations: Comparisons by gender, gender role identity, and socioeconomic status. *Professional School Counseling*, *2*(4), 314. https://www.proquest.com/scholarly-journals/childrens-occupational-aspirations-comparisons/docview/213279779/se-2

Shapiro, L. R. (2009). Eyewitness testimony for a simulated juvenile crime by male and female criminals with consistent or inconsistent gender-role characteristics. *Journal of Applied Developmental Psychology, 30*(6), 649–666. https://doi.org/10.1016/J.APPDEV.2009.07.007

Shinar, E. H. (1975). Sexual stereotypes of occupations. *Journal of Vocational Behavior, 7*(1), 99–111. https://doi.org/10.1016/0001-8791(75)90037-8

Signorella, M. L., Bigler, R. S., & Liben, L. S. (1997). A meta-analysis of children's memories for own-sex and other-sex information. *Journal of Applied Developmental Psychology, 18*(3), 429–445. https://doi.org/10.1016/S0193-3973(97)80009-3

Signorella, M. L., & Liben, L. S. (1984). Recall and reconstruction of gender-related pictures: Effects of attitude, task difficulty, and age. *Child Development, 55*(2), 393. https://doi.org/10.2307/1129951

Solbes-Canales, I., Valverde-Montesino, S., & Herranz-Hernández, P. (2020). Socialization of gender stereotypes related to attributes and professions among young spanish school-aged children. *Frontiers in Psychology, 11*, 609. https://doi.org/10.3389/FPSYG.2020.00609/BIBTEX

Stahlberg, D., Braun, F., Irmen, L., & Sczesny, S. (2007). Representation of the sexes in language. In K. Fiedler (Ed.), *Social communication*. *A volume in the series Frontiers of Social Psychology* (pp. 163–187). Psychology Press.

Stangor, C. (1988). Stereotype accessibility and information processing. *Personality and Social Psychology Bulletin, 14*(4), 694–708. https://doi.org/10.1177/0146167288144005

Stangor, C., & Duan, C. (1991). Effects of multiple task demands upon memory for information about social groups. *Journal of Experimental Social Psychology, 27*(4), 357–378. https://doi.org/10.1016/0022-1031(91)90031-Z

Stangor, C., & McMillan, D. (1992). Memory for expectancy-congruent and expectancy-incongruent information: A review of the social and social developmental literatures. *Psychological Bulletin, 111*(1), 42. https://doi.org/10.1037/0033-2909.111.1.42

Stangor, C., & Ruble, D. N. (1989). Differential influences of gender schemata and gender constancy on children's information processing and behavior. *Social Cognition, 7*(4), 353–372. https://doi.org/10.1521/SOCO.1989.7.4.353

Stout, J. G., & Dasgupta, N. (2011). When he doesn't mean you: Gender-exclusive language as ostracism. *Personality and Social Psychology Bulletin, 37*(6), 757–769. https://doi.org/10.1177/0146167211406434

Streilein, A., Leach, B., Everett, C., & Morgan, P. (2018). Knowing your worth: Salary expectations and gender of matriculating physician assistant students. *The Journal of Physician Assistant Education, 29*(1), 1–6. https://doi.org/10.1097/JPA.0000000000000180

Tobin, D. D., Menon, M., Menon, M., Spatta, B. C., Hodges, E. V. E., & Perry, D. G. (2010). The intrapsychics of gender: A model of self-socialization. *Psychological Review, 117*(2), 601–622. https://doi.org/10.1037/A0018936

Treleaven, C. (2015). Gender, generation, and jobs: differences in gender role ideologies by age and occupation. *MA Research Paper*. https://ir.lib.uwo.ca/sociology_masrp/1

Vervecken, D., Hannover, B., & Wolter, I. (2013). Changing (S) expectations: How gender fair job descriptions impact children's perceptions and interest regarding traditionally male occupations. *Journal of Vocational Behavior, 82*(3), 208–220. https://doi.org/10.1016/j.jvb.2013.01.008

Wasserstein, R. L., & Lazar, N. A. (2016). The ASA statement on *p*-values: Context, process, and purpose. *The American Statistician, 70*(2), 129–133. https://doi.org/10.1080/00031305.2016.1154108

Weisgram, E. S., Bigler, R. S., & Liben, L. S. (2010). Gender, values, and occupational interests among children, adolescents, and adults. *Child Development, 81*(3), 778–796. https://doi.org/10.1111/j.1467-8624.2010.01433.x

White, M. J., & White, G. B. (2006). Implicit and explicit occupational gender stereotypes. *Sex Roles, 55*(3–4), 259–266. https://doi.org/10.1007/s11199-006-9078-z

Young, V. (2011). *The secret thoughts of successful women: Why capable people suffer from the impostor syndrome and how to thrive in spite of it*. Currency.